



# **STIC Search Report**

## **Biotech-Chem Library**

STIC Database Tracking Number: 103481

**To: Laurie Scheiner**  
**Location: CM1-8E12**  
**Art Unit: 1648**  
**Tuesday, September 16, 2003**

**Case Serial Number: 09/821877**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**CM1-1E05**  
**Phone: 308-4994**

**beverly.shears@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

STIC-Biotech/ChemLib

103 481

From: Scheiner, Laurie  
Sent: Tuesday, September 09, 2003 9:04 PM  
To: STIC-Biotech/ChemLib  
Subject: seq. search request

RECEIVED

SEP 10 2003

Please search SEQ ID NO:1 of application S.N. 09/821,877. Thanks!

STIC-BIOTECH DIVISION  
(STIC)

Laurie Scheiner  
Art Unit 1648  
CM1 8E05  
308-1122  
8E12

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**

## SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

### Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

### STAFF USE ONLY

Date completed: <u>09-16-03</u>	Search Site	Vendors
Searcher: <u>Beverly C 4994</u>	_____ STIC	_____ IG
Terminal time: <u>20</u>	_____ CM-1	_____ STN
Elapsed time: _____	_____ Pre-S	_____ Dialog
CPU time: _____	Type of Search	_____ APS
Total time: <u>23</u>	_____ N.A. Sequence	_____ Geninfo
Number of Searches: _____	_____ A.A. Sequence	_____ SDC
Number of Databases: <u>1</u>	_____ Structure	_____ DARC/Questel
	_____ Bibliographic	<input checked="" type="checkbox"/> Other <u>CGN</u>

**This Page Blank (uspto)**

Scheiner, L.  
09/18/21877 Page 1  
Seq. ID 1 w/ Interf

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2003, 01:28:16 ; Search time 2810 Seconds

(without alignments)  
13938.229 Million cell updates/sec

Title: US-09-821-877-1

Perfect score: 1181  
Sequence: 1 atggggcagaaattcttcac.....tacatttaaccctaataaa 1181

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 33363688 seqs, 1658189874 residues.

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_NA\_Main:

1: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

7: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

8: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

9: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

10: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

11: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

12: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

13: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

14: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

15: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

16: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

17: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

18: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

19: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

20: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

21: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

22: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

23: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

24: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

25: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

26: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

27: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

28: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

29: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

30: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

31: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

32: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

33: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

34: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

35: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

36: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

37: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

38: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

39: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

40: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

41: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

42: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

43: /cgn2\_6/ptodata/2/pna/US0928A\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US100A\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US100B\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US101A\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US101B\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US102A\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US102B\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US103A\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US103B\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US104A\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US104B\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
76: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
77: /cgn2\_6/ptodata/2/pna/US6023A\_COMB.seq.\*  
78: /cgn2\_6/ptodata/2/pna/US6023B\_COMB.seq.\*  
79: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
80: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
81: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*  
82: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*  
83: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*  
84: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*  
85: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*  
86: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*  
87: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*  
88: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*  
89: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*  
90: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*  
91: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*  
92: /cgn2\_6/ptodata/2/pna/US6037\_COMB.seq.\*  
93: /cgn2\_6/ptodata/2/pna/US6038\_COMB.seq.\*  
94: /cgn2\_6/ptodata/2/pna/US6039\_COMB.seq.\*  
95: /cgn2\_6/ptodata/2/pna/US6040\_COMB.seq.\*  
96: /cgn2\_6/ptodata/2/pna/US6041\_COMB.seq.\*  
97: /cgn2\_6/ptodata/2/pna/US6042\_COMB.seq.\*  
98: /cgn2\_6/ptodata/2/pna/US6043\_COMB.seq.\*  
99: /cgn2\_6/ptodata/2/pna/US6044\_COMB.seq.\*  
100: /cgn2\_6/ptodata/2/pna/US6045\_COMB.seq.\*  
101: /cgn2\_6/ptodata/2/pna/US6046\_COMB.seq.\*  
102: /cgn2\_6/ptodata/2/pna/US6047\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	1181	100.0	1181	1	PCT-US02-09227-1	Sequence 1, Appli
2	1181	100.0	1181	34	US-09-821-877-1	Sequence 1, Appli
3	1128.2	95.5	8007	35	US-09-837-297-3	Sequence 3, Appli
4	1117.2	94.6	1201	4	US-07-847-951A-218	Sequence 218, App

5	1117.2	94.6	1201	5	US-07-846-318-216
6	1117.2	94.6	1205	4	US-07-847-951A-215
7	1117.2	94.6	1285	5	US-08-036-218-213
8	1020.6	86.4	3248	4	US-10-130-915-918
9	1010.2	85.5	1203	31	US-09-718-095-8
10	1005.2	85.1	1203	1	PCT-US02-33263-45
11	1005.2	85.1	1203	45	US-10-287-994-45
12	1005.2	85.1	1203	52	US-10-410-897-45
13	1005.2	85.1	1203	52	US-10-410-913-45
14	1005.2	85.1	1203	52	US-10-410-930-45
15	1005.2	85.1	1203	52	US-10-410-945-45
16	1005.2	85.1	1203	52	US-10-410-962-45
17	1005.2	85.1	1203	52	US-10-410-980-45
18	1005.2	85.1	1203	52	US-10-410-997-45
19	1005.2	85.1	1203	52	US-10-411-012-45
20	1005.2	85.1	1203	52	US-10-411-026-45
21	1005.2	85.1	1203	52	US-10-411-037-45
22	1005.2	85.1	1203	52	US-10-411-043-45
23	1005.2	85.1	1203	52	US-10-411-044-45
24	1005.2	85.1	1203	52	US-10-411-049-45
25	1001.8	84.8	4084	33	US-09-781-891D-22
26	1001.8	84.8	4496	33	US-09-781-891D-22
27	1000.6	84.7	1200	9	US-08-667-859-5
28	1000.6	84.7	1200	13	US-08-632-981-5
29	997.6	84.5	2566	46	US-10-130-915A-4
30	988.6	83.7	1047	37	US-09-890-755A-4
31	984	83.3	1365	51	US-10-365-620-33
32	984	83.3	2037	51	US-10-365-620-38
33	984	83.3	2037	51	US-10-365-620-31
34	962.2	81.5	1286	1	PCT-US02-13164-3
35	962.2	81.5	1286	1	PCT-US02-13644-1
36	962.2	81.5	1286	46	US-10-132-829-3
37	962.2	81.5	1286	46	US-10-133-907-3
38	962.2	81.5	1286	46	US-10-136-819-1
39	947.4	80.2	1122	31	US-09-718-095-54
40	946.6	80.2	1122	31	US-09-718-095-55
41	936.8	79.3	1104	31	US-09-718-095-57
42	933.4	79.0	1084	31	US-09-718-095-58
43	929.4	78.7	1102	31	US-09-718-095-58
44	923	78.2	1132	31	US-09-718-095-52
45	914.8	77.5	1076	31	US-09-718-095-55

[illegible]

## ALIGNMENTS

6/26/01  
apric

```

RESULT 1
PCT-US02-09227-1
, Sequence 1, Application PC/TUS0209227
, GENERAL INFORMATION:
, APPLICANT: Abbott Laboratories
, APPLICANT: Coleman, Paul F.
, APPLICANT: Mushahwar, Isa K.
, TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
, TITLE OF INVENTION: And Methods of Detection Thereof
, FILE REFERENCE: 6794, US-01
, CURRENT APPLICATION NUMBER: PCT/US02/09227
, CURRENT FILING DATE: 2002-03-26
, NUMBER OF SEQ ID NOS: 8
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 1
, LENGTH: 1161
, TYPE: DNA
, ORGANISM: Hepatitis B Virus
PCT-US02-09227-1

```

Query Match	100.0%	Score 1181;	DB 1;	Length 1181;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1181; Conservative	0;	Mismatches	0;	Gaps 0;

QY 1 ATGGGCGAATCTTTCCACGACAATCTGTGGATTCTTCCGACCACCACTTGGAT 60

Db 1 ATGGGCGAATCTTTCCACGACAATCTGTGGATTCTTCCGACCACCACTTGGAT 60

QY	61	CGACGCTTCAGAGAAACCCAAATCCAGATTTGGGACTTTCATCTCCCAACGAGCAACC	120
Db	61	CCAGCTTCAGAGAAACCCAAATCCAGATTTGGGACTTTCATCTCCCAACGAGCAACC	120
QY	121	TGGCCAGACGCCAACGAGTAGAGCTGGAGCATTTGGACTGGGGTTTCACCCCACGCCAC	180
Db	121	TGGCCAGACGCCAACGAGTAGAGCTGGAGCATTTGGACTGGGGTTTCACCCCACGCCAC	180
QY	181	GGAGGCTTTTGGGGGTGGAGCCCTCAGGCTCAGGGGCTTAACAAACCTTGGCCAGCAAT	240
Db	181	GGAGGCTTTTGGGGGTGGAGCCCTCAGGCTCAGGGGCTTAACAAACCTTGGCCAGCAAT	240
QY	241	CCGCGCTCTGCTTCCACCAATCGCCAGTCAAGGAAGCAGGCTACCCGCTGTCCAACT	300
Db	241	CCGCGCTCTGCTTCCACCAATCGCCAGTCAAGGAAGCAGGCTACCCGCTGTCCAACT	300
QY	301	TTGAGAAACACTCATCTCTCAAGCCATCGAGTGGAACTTCCACCACTTTCACCAACTCTG	360
Db	301	TTGAGAAACACTCATCTCTCAAGCCATCGAGTGGAACTTCCACCACTTTCACCAACTCTG	360
QY	361	CAAGATCCAGAGTAGAGGTCTGTATTTTCCCTGCTGCTGAGTCCAGTTACAGAAACGTA	420
Db	361	CAAGATCCAGAGTAGAGGTCTGTATTTTCCCTGCTGAGTCCAGTTACAGAAACGTA	420
QY	421	AACCGTTTCCGACTACTGTCTCTCCCATATCGTCAATCTTTCGAGGATTTGGGACCT	480
Db	421	AACCGTTTCCGACTACTGTCTCTCCCATATCGTCAATCTTTCGAGGATTTGGGACCT	480
QY	481	GGCGGGAAACATGGAGAAATCATCATTCAGGATTTCTTAGAGCCCGCTCGGTTCACAGCG	540
Db	481	GGCGGGAAACATGGAGAAATCATCATTCAGGATTTCTTAGAGCCCGCTCGGTTCACAGCG	540
QY	541	GGGTTTTTCTGTGACAGAAATCTCTCAATACCGAGAGTCTAGACTCGTGCTGACT	600
Db	541	GGGTTTTTCTGTGACAGAAATCTCTCAATACCGAGAGTCTAGACTCGTGCTGACT	600
QY	601	TCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTTGCAGTCCCAACC	660
Db	601	TCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTTGCAGTCCCAACC	660
QY	661	TCCAAATCACTACCAACCTCGCTGCTCCCAACTGTCTGGTTATTCGCTGGAATGTCTG	720
Db	661	TCCAAATCACTACCAACCTCGCTGCTCCCAACTGTCTGGTTATTCGCTGGAATGTCTG	720
QY	721	CGGGGTTTATCATCTTCTCTTCATCCTGTCTGATGCTCATCTTCTTGTGGTTCTT	780
Db	721	CGGGGTTTATCATCTTCTCTTCATCCTGTCTGATGCTCATCTTCTTGTGGTTCTT	780
QY	781	CTGGAATATCAAGGTATGTGCCCGTTTGTCTTAAATTCAGAGATCTTCAACCAACAGC	840
Db	781	CTGGAATATCAAGGTATGTGCCCGTTTGTCTTAAATTCAGAGATCTTCAACCAACAGC	840
QY	841	ACGGGACCATCAGAGCTGTGACAGATCTGTCTCAAGAAACCTTATGTATCCCTCTGT	900
Db	841	ACGGGACCATCAGAGCTGTGACAGATCTGTCTCAAGAAACCTTATGTATCCCTCTGT	900
QY	901	TGCTGTACAAACCTTGGAGTGAACCTGCACCTGTATTCCTATCCATCATCTCTGGGCT	960
Db	901	TGCTGTACAAACCTTGGAGTGAACCTGCACCTGTATTCCTATCCATCATCTCTGGGCT	960
QY	961	TTTGGAAATTCCTATGGGAGTGGGCGTCAAGCCGTTTCTCTGGCTCAAGTTTACTAGTG	1020
Db	961	TTTGGAAATTCCTATGGGAGTGGGCGTCAAGCCGTTTCTCTGGCTCAAGTTTACTAGTG	1020
QY	1021	CCATTGTGTCAGTGGTCTGTAGAGGCTTTCGCCCACTGTGGCTTTCAGTTATATGAGTG	1080
Db	1021	CCATTGTGTCAGTGGTCTGTAGAGGCTTTCGCCCACTGTGGCTTTCAGTTATATGAGTG	1080
QY	1081	ATGTGTATCTGGGGCCCAAGTCTGTACCAACTTTCAGTCCCTTTTACCGCTGTATACA	1140
Db	1081	ATGTGTATCTGGGGCCCAAGTCTGTACCAACTTTCAGTCCCTTTTACCGCTGTATACA	1140



QY 1141 ATTTCTTTTGTCTTGGGTATATTAACCTTAATAA 1181  
 DB 1141 ATTTCTTTTGTCTTGGGTATATTAACCTTAATAA 1181

## RESULT 2

US-09-821-877-1  
 ; Sequence 1, Application US/09821877  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abbott Laboratories  
 ; APPLICANT: Coleman, Paul F.  
 ; APPLICANT: Mushahwar, Isa K.  
 ; TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant  
 ; TITLE OF INVENTION: And Methods Of Detection Thereof  
 ; FILE REFERENCE: 6794.US.01  
 ; CURRENT APPLICATION NUMBER: US/09/821,877  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1181  
 ; TYPE: DNA  
 ; ORGANISM: Hepatitis B Virus  
 US-09-821-877-1

Query Match 100.0%; Score 1181; DB 34; Length 1181;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGAGAAATCTTTCACAGCAATCTCTGGGATTTCTTCCGACCAAGTTGAT 60  
 DB 1 ATGGGGAGAAATCTTTCACAGCAATCTCTGGGATTTCTTCCGACCAAGTTGAT 60  
 QY 61 CGAGCTTCAGAGCAAAACCAACATCCAGATTGGGAACTTAAATCCCAAGACACC 120  
 DB 61 CGAGCTTCAGAGCAAAACCAACATCCAGATTGGGAACTTAAATCCCAAGACACC 120  
 QY 121 TGGCCAGAGCCCAAGGTAGAGCTGAGCAATTCGAGCTGGGGTTCCACCCACGAC 180  
 DB 121 TGGCCAGAGCCCAAGGTAGAGCTGAGCAATTCGAGCTGGGGTTCCACCCACGAC 180  
 QY 181 GGAGGCTTTTGGGGTGAGGCCCTCAGGCTCAGGGGATTAACAACCTTGCAGCAAT 240  
 DB 181 GGAGGCTTTTGGGGTGAGGCCCTCAGGCTCAGGGGATTAACAACCTTGCAGCAAT 240  
 QY 241 CGGCTCTCTGCTTCCAGCAATGGCAATGCAAGTGAAGGACGCTTACCCGCTGCTCACT 300  
 DB 241 CGGCTCTCTGCTTCCAGCAATGGCAATGCAAGTGAAGGACGCTTACCCGCTGCTCACT 300  
 QY 301 TTGAGAAACACTCATCTCTCAAGCCATGAGTGAAGTCCAGCAACTTTCACCAAACTCTG 360  
 DB 301 TTGAGAAACACTCATCTCTCAAGCCATGAGTGAAGTCCAGCAACTTTCACCAAACTCTG 360  
 QY 361 CAAGATCCCAAGAGTAGAGTGTATTTCCCTGTGTGTGTCCAGTTCAAGAAACAGTA 420  
 DB 361 CAAGATCCCAAGAGTAGAGTGTATTTCCCTGTGTGTGTCCAGTTCAAGAAACAGTA 420  
 QY 421 AACCTGTTCGCACTACTGTCTCTCCCATATGTCATTTTTCAGAGATTGGGACCTT 480  
 DB 421 AACCTGTTCGCACTACTGTCTCTCCCATATGTCATTTTTCAGAGATTGGGACCTT 480  
 QY 481 GGGCGAATCATGAGAAACATCATCAGGATTCCTAGGACCCCTGTGTGTTCAGAGG 540  
 DB 481 GGGCGAATCATGAGAAACATCATCAGGATTCCTAGGACCCCTGTGTGTTCAGAGG 540  
 QY 541 GGGTTTCTTGTGTGAACAAGATCTCAAAATACCGCAGAGTCTAGACTGTGTGTGACT 600  
 DB 541 GGGTTTCTTGTGTGAACAAGATCTCAAAATACCGCAGAGTCTAGACTGTGTGTGACT 600  
 QY 601 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTGTCTTGGCAAAATTCGAGTCCCAACC 660  
 DB 601 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTGTCTTGGCAAAATTCGAGTCCCAACC 660

QY 661 TCCAATCACTACCAACCTCTGCTCTCAACTGTGCTCTGTATATGCTGATGCTG 720  
 DB 661 TCCAATCACTACCAACCTCTGCTCTCAACTGTGCTCTGTATATGCTGATGCTG 720  
 QY 721 CGGCTTTTATCATTTCTCTTTCATCTGCTGCTATGCTCATTTCTTGTGTTCTT 780  
 DB 721 CGGCTTTTATCATTTCTCTTTCATCTGCTGCTATGCTCATTTCTTGTGTTCTT 780  
 QY 781 CTGACATTAAGGATGATGTCGGTGTGCTCTCAATTCAGATCTTCAACCAACGAC 840  
 DB 781 CTGACATTAAGGATGATGTCGGTGTGCTCTCAATTCAGATCTTCAACCAACGAC 840  
 QY 841 ACGGACCATGAGAGCCCTGACAGACTCTCTGCTCAAGAACCTCATATCCCTCTGT 900  
 DB 841 ACGGACCATGAGAGCCCTGACAGACTCTCTGCTCAAGAACCTCATATCCCTCTGT 900  
 QY 901 TGCTGTACAAAACCTTGGATGGAACCTGACCTGTATTTCCATTCATCTGAGCT 960  
 DB 901 TGCTGTACAAAACCTTGGATGGAACCTGACCTGTATTTCCATTCATCTGAGCT 960  
 QY 961 TTGGAAATTCCTATGAGAGTGGGCTCAGGCCGTTCTCTGCTGCTAGTTACTAGT 1020  
 DB 961 TTGGAAATTCCTATGAGAGTGGGCTCAGGCCGTTCTCTGCTGCTAGTTACTAGT 1020  
 QY 1021 CCATTGTTCAAGTGTGTGATGAGGCTTTCGCCCACTGTTTGGCTTCAATTATGATG 1080  
 DB 1021 CCATTGTTCAAGTGTGTGATGAGGCTTTCGCCCACTGTTTGGCTTCAATTATGATG 1080  
 QY 1081 ATGTTGACTGGGGGCCAAGTCTGTACACATCTTGAATCTCTTTTACCGCTGTACCA 1140  
 DB 1081 ATGTTGACTGGGGGCCAAGTCTGTACACATCTTGAATCTCTTTTACCGCTGTACCA 1140  
 QY 1141 ATTTCTTTTGTCTTGGGTATATTAACCTTAATAA 1181  
 DB 1141 ATTTCTTTTGTCTTGGGTATATTAACCTTAATAA 1181

## RESULT 3

US-09-837-297-3  
 ; Sequence 3, Application US/09837297  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RYU, WANG SHICK  
 ; TITLE OF INVENTION: Hepatitis B virus vectors for gene therapy  
 ; FILE REFERENCE:  
 ; CURRENT APPLICATION NUMBER: US/09/837,297  
 ; PRIOR APPLICATION NUMBER: KR2000-21070  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: Koparentin 1.71  
 ; SEQ ID NO 3  
 ; LENGTH: 8007  
 ; TYPE: DNA  
 ; ORGANISM: HBV  
 ; FEATURE:  
 ; NAME/KEY: gene  
 ; LOCATION: (1)..(8007)  
 ; OTHER INFORMATION: Prototype vector of HBV  
 US-09-837-297-3

Query Match 95.5%; Score 1128.2; DB 35; Length 8007;  
 Best Local Similarity 97.2%; Pred. No. 0;  
 Matches 1148; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGGGGAGAAATCTTTCACAGCAATCTCTGGGATTTCTTCCGACCAAGTTGAT 60  
 DB 1021 ATGGGGAGAAATCTTTCACAGCAATCTCTGGGATTTCTTCCGACCAAGTTGAT 1090  
 QY 61 CGAGCTTCAGAGCAAAACCAACATCCAGATTGGGAACTTAAATCCCAAGACACC 120  
 DB 1091 CGAGCTTCAGAGCAAAACCAACATCCAGATTGGGAACTTAAATCCCAAGACACC 1150  
 QY 121 TGGCCAGAGCCCAAGGTAGAGCTGAGCAATTCGAGCTGGGGTTCCACCCACGAC 180

Db 1151 TGCCAGACGCCAACAAGTAGAGCTGGAGCAATGGGGCTGGGTTTCACCCCACGGCAC 1210  
QY 181 GAGAGCCTTTTGGGGTGGAGACCTCTCAGGCTCAGGGGCTAACACAAACCTTGGCCAGCAAT 240  
Db 1211 GAGAGCCTTTTGGGGTGGAGACCTCTCAGGCTCAGGGGCTAACACAAACCTTGGCCAGCAAT 1270  
QY 241 CCGCTCTCTGCTTTCACCAATCGCAGCTCAGGAAGCAGCTACCCGCTGTCTCCACT 300  
Db 1271 CCGCTCTCTGCTTTCACCAATCGCAGCTCAGGAAGCAGCTACCCGCTGTCTCCACT 1330  
QY 301 TTGAGAAACACTCATCTCTCAAGCCATCAGTGGAACTCCAACTTTCACCAACTCTG 360  
Db 1331 TTGAGAAACACTCATCTCTCAGGCCATCAGTGGAACTCCAACTTTCACCAACTCTG 1390  
QY 361 CAAGATCCCGAGTGGAGGCTGTATTTCCCTGCTGGTGGCTCCACTTTCAGGAACAGTA 420  
Db 1391 CAAGATCCCGAGTGGAGGCTGTATTTCCCTGCTGGTGGCTCCACTTTCAGGAACAGTA 1450  
QY 421 AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTTCAGAGATTGGGAGCCT 480  
Db 1451 AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTTCAGAGATTGGGAGCCT 1510  
QY 481 GCGCGAACAATGAGAAACATCATCATCAGATTCTCAGAACCTCTGTGTATCAGGCG 540  
Db 1511 GCGCTGAACATGAGAAACATCATCATCAGATTCTCAGAACCTCTGTGTATCAGGCG 1570  
QY 541 GGGTTTTTCTGTGTGACAAATTCCTCAATACCGAGAGTCTAGAGCTGTGGTGGACT 600  
Db 1571 GGGTTTTTCTGTGTGACAAATTCCTCAATACCGAGAGTCTAGAGCTGTGGTGGACT 1630  
QY 601 TCTCTCAATTTCTAGGGGAACTACCGTGTGTCTTGGCCAAATTTGCGAGTCCCAACC 660  
Db 1631 TCTCTCAATTTCTAGGGGAACTACCGTGTGTCTTGGCCAAATTTGCGAGTCCCAACC 1690  
QY 661 TCCAAATCTCACCACCTCTCTGTCTTCCAACTTGTCTGTGTATTCGCTGAGTGTCTG 720  
Db 1691 TCCAAATCTCACCACCTCTCTGTCTTCCAACTTGTCTGTGTATTCGCTGAGTGTCTG 1750  
QY 721 GCGGTTTTTATCATCTCTCTCTCATATCTGTGTATGCTCATCTCTGTGTGTCTT 780  
Db 1751 GCGGTTTTTATCATCTCTCTCTCATATCTGTGTATGCTCATCTCTGTGTGTCTT 1810  
QY 781 CTGGAATCATAGGATGTGGCCGTTTGTCTCTAATTCAGAGATCTTCAACCCAGC 840  
Db 1811 CTGGAATCATAGGATGTGGCCGTTTGTCTCTAATTCAGAGATCTTCAACCCAGC 1870  
QY 841 ACGGACCATGACAGGCTGACGCTCTGTCTCAGGAACCTTATGTATCCCTCTGT 900  
Db 1871 ACGGACCATGACGAGCTGACGCTCTGTCTCAGGAACCTTATGTATCCCTCTGT 1930  
QY 901 TGCGTACAAACCTTGGAGTGGAACTGACCTGTATTCCTCATCCCATATCCTGGGCT 960  
Db 1931 TGCGTACAAACCTTGGAGTGGAACTGACCTGTATTCCTCATCCCATATCCTGGGCT 1990  
QY 961 TTGGAATAATTCATATGGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTTACTAGTG 1020  
Db 1991 TTGGAATAATTCATATGGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTTACTAGTG 2050  
QY 1021 CCATTTTTCAGTGGTTCGATAGGGCTTCCCCACTGTGTGGCTTCAAGTTATGTAGTG 1080  
Db 2051 CCATTTTTCAGTGGTTCGATAGGGCTTCCCCACTGTGTGGCTTCAAGTTATGTAGTG 2110  
QY 1081 ATGTGTACTGGGGGCAAGCTGTATACCACTGTGAGTCCCTTTTACCGGCTTACCA 1140  
Db 2111 ATGTGTACTGGGGGCAAGCTGTATACCACTGTGAGTCCCTTTTACCGGCTTACCA 2170  
QY 1141 ATTTCTTTTGTCTTGGGTATACATTTAAACCTTAATAA 1181  
Db 2171 ATTTCTTTTGTCTTGGGTATACATTTAAACCTTAATAA 2211

RESULT 4

US-07-847-951A-218  
; Sequence 218, Application US/07847951A  
; GENERAL INFORMATION:  
; APPLICANT: Paolletti, Enzo  
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
; TITLE OF INVENTION: STRAIN  
; NUMBER OF SEQUENCES: 462  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; ADDRESSEE: c/o William S. Frommer  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/847,951A  
; FILING DATE: 19920326  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 218:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1201 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-07-847-951A-218  
  
Query Match 94.6%; Score 1117.2; DB 4; Length 1201;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 113; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
  
QY 1 ATGGGCGAGAAATTTTTCACACAGCAATCTCTGGGATTTCTTCCGACACAGTTGGAT 60  
Db 32 ATGGGCGAGAAATTTTTCACACAGCAATCTCTGGGATTTCTTCCGACACAGTTGGAT 91  
QY 61 CCAAGCTTCAGAGCAACACCAACATTCAGATTGGGACTTCAATCCCAAGACACC 120  
Db 92 CCAAGCTTCAGAGCAACACCAACATTCAGATTGGGACTTCAATCCCAAGACACC 151  
QY 121 TGCGCAGAGCGCAACAAGTAGGAGCTGGAGCATTTGGGATCTGAGGCTTACCCGCGCAC 180  
Db 152 TGCGCAGAGCGCAACAAGTAGGAGCTGGAGCATTTGGGATCTGAGGCTTACCCGCGCAC 211  
QY 181 GAGAGCCTTTTGGGGTGGAGCCTCTCAGGCTCAGGGGATTAACAACAACTTGGCCAGCAAT 240  
Db 212 GAGAGCCTTTTGGGGTGGAGCCTCTCAGGCTCAGGGGATTAACAACAACTTGGCCAGCAAT 271  
QY 241 CCGCTCTCTGCTTTCACCAATCGCAGCTCAGGAAGCAGCTACCCGCTGTCTCCACT 300  
Db 272 CCGCTCTCTGCTTTCACCAATCGCAGCTCAGGAAGCAGCTACCCGCTGTCTCCACT 331  
QY 301 TTGAGAAACACTCATCTCTCAAGCCATCAGTGGAACTCCCAACTTTCACCAACTCTG 360  
Db 332 TTGAGAAACACTCATCTCTCAAGCCATCAGTGGAACTCCCAACTTTCACCAACTCTG 391  
QY 361 CAAGATCCCGAGTGGAGGCTGTATTTCCCTGCTGGTGGCTCCACTTTCAGGAACAGTA 420  
Db 392 CAAGATCCCGAGTGGAGGCTGTATTTCCCTGCTGGTGGCTCCACTTTCAGGAACAGTA 451  
QY 421 AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTTCAGAGATTGGGAGCCT 480

```

Db 452 AACCTGTCGACTACTGCTCTCCCTTATGTCATCTTCTGAGATGGGGACCT 511
Qy 481 GCGGGAACATGAGAACATCATCAGGATTCCTAGGACCCCTGCTGCTTACAGGCG 540
Db 512 GCGGGAACATGAGAACATCATCAGGATTCCTAGGACCCCTGCTGCTTACAGGCG 571
Qy 541 GGGTTTCTTGTGACAGAAATCTCACAATACCGAGATCTAGACTGTGTGACT 600
Db 572 GGGTTCTTGTGACAGAAATCTCACAATACCGAGATCTAGACTGTGTGACT 631
Qy 601 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660
Db 632 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 691
Qy 661 TCCATCATCTACCAACCTCTGCTCTCAACTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 692 TCCATCATCTACCAACCTCTGCTCTCAACTGCTGCTGCTGCTGCTGCTGCTGCT 751
Qy 721 GCGGTTTATCATCTTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 752 GCGGTTTATCATCTTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811
Qy 781 CTGACATCAAGATGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 812 CTGACATCAAGATGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871
Qy 841 AGGGGACATGAGAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 872 AGGGGACATGAGAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931
Qy 901 TCTGTACAAACCTTCGAGTGAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 932 TCTGTACAAACCTTCGAGTGAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
Qy 961 TTCGAAAAATTCCTATGAGAGTGGGCTGAGCCGCTTCTGCTGCTGCTGCTGCT 1020
Db 992 TTCGAAAAATTCCTATGAGAGTGGGCTGAGCCGCTTCTGCTGCTGCTGCTGCT 1051
Qy 1021 CCATTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1052 CCATTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1111
Qy 1081 ATGTTGATGAGGAGCAAGTCTGATACCATCTTGAAGTCCCTTTTACCGCTTACCA 1140
Db 1112 ATGTTGATGAGGAGCAAGTCTGATACCATCTTGAAGTCCCTTTTACCGCTTACCA 1171
Qy 1141 ATTTTCTTTTGTCTTGGGTATACATTTAA 1170
Db 1172 ATTTTCTTTTGTCTTGGGTATACATTTAA 1201

```

RESULT 5  
US-08-036-218-216  
Sequence 216, Application US/08036218

GENERAL INFORMATION:

APPLICANT: Paolletti, Enzo  
APPLICANT: Perkus, Marion E.  
APPLICANT: Taylor, Jill  
APPLICANT: Tartaglia, James  
APPLICANT: Norton, Elizabeth K.  
APPLICANT: Riviere, Michel  
APPLICANT: de Taisne, Charles  
APPLICANT: Limbach, Keith J.  
APPLICANT: Johnson, Gerard P.  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE STRAIN  
NUMBER OF SEQUENCES: 221  
CORRESPONDENCE ADDRESS:  
ADDRESSES: William S. Frommer c/o  
ADDRESSES: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York

```

STATE: New York
COUNTRY: United State of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,218
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/713,967
FILING DATE: 11-JUN-1991
APPLICATION NUMBER: US 07/666,056
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2291
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 1201 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

US-08-036-218-216

Query Match 94.6%; Score 1117.2; DB 5; Length 1201;

Best Local Similarity 97.2%; Pred. No. 0; Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```

Qy 1 ATGGGGAGAAATCTTCCACGACAAATCCTCTGGGATTCCTTCCGACCAAGTTGAT 60
Db 32 ATGGGGAGAAATCTTCCACGACAAATCCTCTGGGATTCCTTCCGACCAAGTTGAT 91
Qy 61 CCAGCTTCAGAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGAGACACC 120
Db 92 CCAGCTTCAGAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGAGACACC 151
Qy 121 TGGCCAGAGCCCAACAGATGAGAGCTGAGATTCGAGATGGGGTTACCCCAAGC 180
Db 152 TGGCCAGAGCCCAACAGATGAGAGCTGAGATTCGAGATGGGGTTACCCCAAGC 211
Qy 181 GAGAGCTTTGGGGTGGAGCCCTCAGGCTCAGGACATTAACAACAACCTTGCAGCAAT 240
Db 212 GAGAGCTTTGGGGTGGAGCCCTCAGGCTCAGGACATTAACAACAACCTTGCAGCAAT 271
Qy 241 CGGCTCTGCTTCCACCAATGCGAGTCCAGAGAGGAGCCCTGCTGCTCAGCT 300
Db 272 CGGCTCTGCTTCCACCAATGCGAGTCCAGAGAGGAGCCCTGCTGCTCAGCT 331
Qy 301 TTGAGAAACATCTCATCTTCAAGCCATGAGTGAATCTCACAACCTTTCCACCAACTCTG 360
Db 332 TTGAGAAACATCTCATCTTCAAGCCATGAGTGAATCTCACAACCTTTCCACCAACTCTG 391
Qy 361 CAAGATCCAGAGTGAAGGTCTGATTTCCCTGCTGAGTCCAGATTCAGAGACAGTA 420
Db 392 CAAGATCCAGAGTGAAGGTCTGATTTCCCTGCTGAGTCCAGATTCAGAGACAGTA 451
Qy 421 AACCTGTTCCGACTACTGCTCTCCCATTCATCTTCTCGAGATTGGGAGCCCT 480
Db 452 AACCTGTTCCGACTACTGCTCTCCCATTCATCTTCTCGAGATTGGGAGCCCT 511
Qy 481 GCGGGAACATGAGAACATCATCAGGATTCCTAGGACCCCTGCTGCTTACAGGCG 540
Db 512 GCGGGAACATGAGAACATCATCAGGATTCCTAGGACCCCTGCTGCTTACAGGCG 571
Qy 541 GGGTTTCTTGTGACAGAAATCTCACAATACCGAGATCTAGACTGTGTGACT 600

```

081036,217

5364773

|||||  
Db 572 GGGTCTCTTGTGACAAATCTCTCAATACCGAGAGTCTAGACTCGTGTGACT 631  
Qy 601 TCTCTAAATTTCTAGGGGGAATCAACGCTGTCTTGGCCAAATTTGCGAGTCCCAAC 660  
Db 632 TCTCTCAATTTCTAGGGGGAATCAACGCTGTCTTGGCCAAATTTGCGAGTCCCAAC 691  
Qy 661 TCCAACTCACTCAACCACTCTCTCTCTCAACTGTCTGTATCGTGGATGTGTG 720  
Db 692 TCCAACTCACTCAACCACTCTCTCTCTCAACTGTCTGTATCGTGGATGTGTG 751  
Qy 721 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
Db 752 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 811  
Qy 781 CTGAGACTATCAAGATAGTGTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 840  
Db 812 CTGAGACTATCAAGATAGTGTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 871  
Qy 841 ACGGACATGACAGAGCTCTGACGACTCTGCTCAAGAACTCTATGTATCCCTCTGT 900  
Db 872 ACGGACATGACAGAGCTCTGACGACTCTGCTCAAGAACTCTATGTATCCCTCTGT 931  
Qy 901 TCGCTGACAAACCTTCTGATGAACTGAACTGAACTGAACTGAACTGAACTGAACT 960  
Db 932 TCGCTGACAAACCTTCTGATGAACTGAACTGAACTGAACTGAACTGAACTGAACT 991  
Qy 961 TTGGGAAATTTCTATGAGAGTGGGAGCTCAAGCCGTTCTCCGAGCTCAATTAAGT 1020  
Db 992 TTGGGAAATTTCTATGAGAGTGGGAGCTCAAGCCGTTCTCCGAGCTCAATTAAGT 1051  
Qy 1021 CCATTTGTTGATGATGTTCTGATGAGGCTTTCCGCCACTGTTTGAGTTGATTAATGAGT 1080  
Db 1052 CCATTTGTTGATGATGTTCTGATGAGGCTTTCCGCCACTGTTTGAGTTGATTAATGAGT 1111  
Qy 1081 ATGTTGATCTGGGGGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTTACCA 1140  
Db 1112 ATGTTGATCTGGGGGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTTACCA 1171  
Qy 1141 ATTTCTTTGTTGTTGGGTTATCATTTAA 1170  
Db 1172 ATTTCTTTGTTGTTGGGTTATCATTTAA 1201

RESULT 6  
US-07-847-951A-215  
Sequence 215, Application US/07847951A  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
TITLE OF INVENTION: STRAIN  
NUMBER OF SEQUENCES: 462  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESS: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/847,951A  
FILING DATE: 19920326  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2400

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEO ID NO: 215:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1285 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-847-951A-215

Query Match 94.6%; Score 1117.2; DB 4; Length 1285;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ATGGGGCAGAAATTTTCCACACGATTCCTCTGGGATTCCTTCCGACACAGTTGGAT 60  
Db 116 ATGGGGCAGAAATTTTCCACACGATTCCTCTGGGATTCCTTCCGACACAGTTGGAT 175  
Qy 61 CCAGCTTCAGAGCAAAACCAACCAATTCAGATTGGGACTTCATCCCAAGAGACACC 120  
Db 176 CCAGCTTCAGAGCAAAACCAACCAATTCAGATTGGGACTTCATCCCAAGAGACACC 235  
Qy 121 TGGCCAGAGCGCCAAAGATGAGAGTGGAGCATTTGGAGTGGGTTTACCCACCGCAC 180  
Db 236 TGGCCAGAGCGCCAAAGATGAGAGTGGAGCATTTGGAGTGGGTTTACCCACCGCAC 295  
Qy 181 GAGAGGCTTTTGGGGTGGAGGCTTCAGAGCTCAGGGCATTAACAACCTTCCAGCAAT 240  
Db 296 GAGAGGCTTTTGGGGTGGAGGCTTCAGAGCTCAGGGCATTAACAACCTTCCAGCAAT 355  
Qy 241 CGCCTCTCTGCTTCCACCAATCGCAGTCAAGAGGAGAGCTTACCCGCTGTCTCCACT 300  
Db 356 CGCCTCTCTGCTTCCACCAATCGCAGTCAAGAGGAGAGCTTACCCGCTGTCTCCACT 415  
Qy 301 TTGAGAAACATATCTCTCAAGCATGCAAGTGAATCTCCAACTTTCCACCAACTCTG 360  
Db 416 TTGAGAAACATATCTCTCAAGCATGCAAGTGAATCTCCAACTTTCCACCAACTCTG 475  
Qy 361 CAAGATCCCGAGTGAAGGCTGTATTTCCCTGCTGGTGGCTCCAGTTGAGGACAGTA 420  
Db 476 CAAGATCCCGAGTGAAGGCTGTATTTCCCTGCTGGTGGCTCCAGTTGAGGACAGTA 535  
Qy 421 AACCTGTTCCGACT 480  
Db 536 AACCTGTTCCGACT 595  
Qy 481 GCGCGGAACATGGAAGAACATCAGATGAGATTCTTGAAGCCCTGCTGTGTTACAGGGC 540  
Db 596 GCGCTGAACATGGAAGAACATCAGATGAGATTCTTGAAGCCCTGCTGTGTTACAGGGC 655  
Qy 541 GGGTTTTTCTGTTGAAGAAATCTCAAAATCCGAGAGTCTAGACTGCTGGTGGACT 600  
Db 656 GGGTTTTTCTGTTGAAGAAATCTCAAAATCCGAGAGTCTAGACTGCTGGTGGACT 715  
Qy 601 TCTCTCAATTTCTAGGGGGAATCAACGCTGTCTTGGCCAAATTTGCGAGTCCCAAC 660  
Db 716 TCTCTCAATTTCTAGGGGGAATCAACGCTGTCTTGGCCAAATTTGCGAGTCCCAAC 775  
Qy 661 TCCAACTCACTCAACCACTCTCTCTCTCAACTGTCTGTATCGTGGATGTGTG 720  
Db 776 TCCAACTCACTCAACCACTCTCTCTCTCAACTGTCTGTATCGTGGATGTGTG 835  
Qy 721 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
Db 836 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 895  
Qy 781 CTGAGACTATCAAGATAGTGTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 840  
Db 896 CTGAGACTATCAAGATAGTGTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 955  
Qy 841 ACGGACATGACAGAGCTCTGACGACTCTGCTCAAGAACTCTATGTATCCCTCTCTGT 900

Db 956 ACGGACCATGCCGACCTGATGACTGCTCAAGAACTTAATGATCCCTCTGT 1015  
QY 901 TCGTGTACAAAACCTTCGGATGAAACTGACCTGATATCCATCCATCTGGGCT 960  
Db 1016 TCGTGTACAAAACCTTCGGATGAAACTGACCTGATATCCATCCATCTGGGCT 1075  
QY 961 TTGGGAAAATTCCTATGAGAGTGGGCTCAGCCGCTTCTCGTGGCTCACTAGTG 1020  
Db 1076 TTGGGAAAATTCCTATGAGAGTGGGCTCAGCCGCTTCTCGTGGCTCACTAGTG 1135  
QY 1021 CCATTTGTTCAGTGTTCGTAGGCTTTCCCACTGTTGGCTTCACTATATGATG 1080  
Db 1136 CCATTTGTTCAGTGTTCGTAGGCTTTCCCACTGTTGGCTTCACTATATGATG 1195  
QY 1081 ATGTGTACTGGGGGCAAGCTGTATACACCATTTGAGTCCCTTTTACCGCTTACCA 1140  
Db 1196 ATGTGTATTTGGGGGCAAGCTGTATACACCATTTGAGTCCCTTTTACCGCTTACCA 1255  
QY 1141 ATTTTCTTTTGTCTTTGGTATATATTTAA 1170  
Db 1256 ATTTTCTTTTGTCTTTGGTATATATTTAA 1285

## RESULT 7

US-08-036-218-213

Sequence 213, Application US/08036218

GENERAL INFORMATION:

APPLICANT: Paolletti, Enzo

APPLICANT: Perkus, Marion E.

APPLICANT: Taylor, Jill

APPLICANT: Tartaglia, James

APPLICANT: Norton, Elizabeth K.

APPLICANT: Riviere, Michel

APPLICANT: Limbach, Keith J.

APPLICANT: Johnson, Gerard P.

APPLICANT: Pincus, Steven E.

APPLICANT: Cox, William I.

TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE STRAIN

NUMBER OF SEQUENCES: 221

CORRESPONDENCE ADDRESS:

ADDRESSEE: William S. Frommer C/O

ADDRESSEE: Curtis, Morris &amp; Safford, P.C.

STREET: 530 Fifth Avenue

CITY: New York

STATE: New York

COUNTRY: United State of America

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/036,218

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/713,967

FILING DATE: 11-JUN-1991

APPLICATION NUMBER: US 07/666,056

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2291

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 213:

SEQUENCE CHARACTERISTICS:

LENGTH: 1285 base pairs

TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-036-218-213

Query Match 94.6%; Score 1117.2; DB 5; Length 1285;

Best Local Similarity 97.2%; Pred. No. 0; Mismatches 33; Indels 0; Gaps 0;

Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGGGCGAAGATCTTCCACAGCAATCTCTGGGATCTTCCGACACAGTTGAT 60  
Db 116 ATGGGCGAAGATCTTCCACAGCAATCTCTGGGATCTTCCGACACAGTTGAT 175  
QY 61 CCAGCTTCAGAGCAACCAACCAATCAATTTGGACTTCAATCCCAAGAGACAC 120  
Db 176 CCAGCTTCAGAGCAACCAACCAATCAATTTGGACTTCAATCCCAAGAGACAC 235  
QY 121 TGGCCAGAGCCCAAGAGTGAAGCTGAGAGATTGGACTGGGGTTCAACCCAGC 180  
Db 236 TGGCCAGAGCCCAAGAGTGAAGCTGAGAGATTGGACTGGGGTTCAACCCAGC 295  
QY 181 GGAGGCTTTGGGGTGGAGCCCTCAGGCTCAGGCAATACACAACTTGGACGAAT 240  
Db 296 GGAGGCTTTGGGGTGGAGCCCTCAGGCTCAGGCAATACACAACTTGGACGAAT 355  
QY 241 CCGGCTCTGCTCCACCAATGCGCAGTCAAGAGAGGACCTACCCGCTGTCCACT 300  
Db 356 CCGGCTCTGCTCCACCAATGCGCAGTCAAGAGAGGACCTACCCGCTGTCCACT 415  
QY 301 TTGAGAAACATCTCATCTCAAGCCATGCAATGCAATCTCAACACTTTCACCA 360  
Db 416 TTGAGAAACATCTCATCTCAAGCCATGCAATGCAATCTCAACACTTTCACCA 475  
QY 361 CAAGATCCCAAGAGAGAGTGTGATTTCCCTGCTGGTGTCCAGTTCAAGAGAG 420  
Db 476 CAAGATCCCAAGAGAGAGTGTGATTTCCCTGCTGGTGTCCAGTTCAAGAGAG 535  
QY 421 AACCTGTTCGACTACTGTCTCTCCATATGTCATCTTCTCGAGATTGGAGCCT 480  
Db 536 AACCTGTTCGACTACTGTCTCTCCATATGTCATCTTCTCGAGATTGGAGCCT 595  
QY 481 GCGGGAACATGAGAACATCATCAGATTCTTAGAACCCCTGCTGTGTACAGGCG 540  
Db 596 GCGGGAACATGAGAACATCATCAGATTCTTAGAACCCCTGCTGTGTACAGGCG 655  
QY 541 GGGTTTCTTGTGACAAGAACTCTCAATATCCGAGAGTCTAGTGTGTGCT 600  
Db 656 GGGTTTCTTGTGACAAGAACTCTCAATATCCGAGAGTCTAGTGTGTGCT 715  
QY 601 TCTCTCAATTTCTAGGGGGAACCTACCGTGTCTTGGCCAAATTTCCAGTCCCA 660  
Db 716 TCTCTCAATTTCTAGGGGGAACCTACCGTGTCTTGGCCAAATTTCCAGTCCCA 775  
QY 716 TCCATCACTACCAACCTCTGTCCTCAACTTGTCTGTTATCGCTGATGTGCTG 835  
Db 836 TCCATCACTACCAACCTCTGTCCTCAACTTGTCTGTTATCGCTGATGTGCTG 895  
QY 836 CGGCGTTTATCATCTTCTCTTCAATCTCTGCTGATGCTCATCTTCTTGTGCT 895  
Db 896 CTGACTATCAAGGATGATGTCGCTGTCTCTCAATTTCCAGGATCTTCAACGAC 955  
QY 901 ACAGGACCATGAGAGCTGACGATCTCTCTCAAGAACCTTATGATCTCTCTGT 900  
Db 956 ACAGGACCATGAGAGCTGACGATCTCTCTCAAGAACCTTATGATCTCTCTGT 1015  
QY 901 TCGTGTACAAAACCTTCGGATGAAACTGACCTGATATCCATCCATCTGGGCT 960  
Db 1016 TCGTGTACAAAACCTTCGGATGAAACTGACCTGATATCCATCCATCTGGGCT 1075  
QY 961 TTGGGAAAATTCCTATGAGAGTGGGCTCAGCCGCTTCTCGTGGCTCACTAGTG 1020

```
Db 1076 TTGGGAAATTCCTATGGAGTGGGCTCAGCCGTTTCTCGGCTCAGTTTACTAGT 1135
Qy 1021 CCATTGTTAGAGTTCGAGAGGCTTCCGCCACGTGTTGGCTTCAATTATATGATG 1080
Db 1136 CCATTGTTAGAGTTCGAGAGGCTTCCGCCACGTGTTGGCTTCAATTATATGATG 1195
Qy 1081 ATGTTGACTGGGGCCCAAGTCTGTATACCATCTTGAAGTCCCTTTTACCGCTGTACCA 1140
Db 1196 ATGTTGATTTGGGGCCCAAGTCTGTATACCATCTTGAAGTCCCTTTTACCGCTGTACCA 1255
Qy 1141 ATTTTCTTTTGTCTTTGGGTATACATTAA 1170
Db 1256 ATTTTCTTTTGTCTTTGGGTATACATTAA 1285

RESULT 8
US-10-130-915-1
; Sequence 1, Application US/10130915
; GENERAL INFORMATION:
; APPLICANT: Stuyver, Lieven
; APPLICANT: Schinazi, Raymond
; APPLICANT: De Gendt, Sija
; APPLICANT: Van Geyt, Carolina
; APPLICANT: Zoulim, Fabien
; APPLICANT: Fried, Michael
; APPLICANT: Rosseau, Rudi
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR A NEW
; TITLE OF INVENTION: GENOTYPE OF HEPATITIS B VIRUS
; FILE REFERENCE: 16117.000102
; CURRENT APPLICATION NUMBER: US/10/130.915
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: PCT/US00/32108
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/167,206
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3248
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Note:
; OTHER INFORMATION: Artificial Sequence = synthetic construct
US-10-130-915-1

Query Match 86.4%; Score 1020.6; DB 46; Length 3248;
Best Local Similarity 91.6%; Pred. No. 4.8e-289;
Matches 1080; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 3 GGGGCAAGATCTTCCACAGCATCTCTGGGATCTTCCGCCAGCCAGTGGATCC 62
Db 1103 GGGGAAAGACCTTTCGCCACAGATCTCTAGAGATCTTCCGCCAGTGGATCC 1162
Qy 63 AGCCTTCAAGCAAAACCAACATCCAGATTGGAGTTCATCCCAACAGACACTG 122
Db 1163 AGATTCAGAGCAAAATCCAAATCCAGATTGGAGTTCATCCCAAAAGACCTTG 1222
Qy 123 GCCAGAGCCCAACAGTAGAGCTGAGCATTTGGAGCTGGGGTTTCAACCCGACGG 182
Db 1223 GCCAGAGCCCAACAGTAGAGCTGAGCATTTGGAGCTGGGGTTTCAACCCGACGG 1282
Qy 183 AGGCTTTTGGGGGAGGAGCCTCAGGCTCAGGGGATACCAAACTTTGCAGCAATCC 242
Db 1283 AGGCTTTTGGGGGAGGAGCCTCAGTCTCAGGGGACACTTAACAATTGGCCAGATCC 1342
Qy 243 GCCTCTGCTTCCACCAATCCGAGTCAAGAAAGCAGGCTTACCCGCTGTCTCCACTTT 302
Db 1343 GCCTCTGCTTCCACCAATCCGAGTCAAGAAAGCAGGCTTACCCGCTGTCTCCACTTT 1402
Qy 303 GAGAAACACTATCTTCAAGCATGAGTGAATCTCAACAATTTCACCAACTCTGCA 362
Db 303 GAGAAACACTATCTTCAAGCATGAGTGAATCTCAACAATTTCACCAACTCTGCA 362
```

```
Db 1403 AAGAGACAGTATCTCTCAGGCCATGCAATGGAACCTTACAGCATTCACCAAGCTTTACA 1462
Qy 363 AGATCCCAAGTAGAGAGTCTGATTTCCCTGCTGAGCTCCAGTTCAAGAACTAA 422
Db 1463 AATCCCAAGTAGAGAGCTGATTTCCCTGCTGAGCTCCAGTTCAAGAACTAA 1522
Qy 423 CCTGTTCCGACTACTGTCTCTCCATATGCTCAATCTTTCTGAGAGATTGGGACCTTCG 482
Db 1523 CCTGTTCCGACTACTGTCTCTCCATATGCTCAATCTTTCTGAGAGATTGGGACCTTCG 1582
Qy 483 GGGGAACATGAGAAATACATACATCAGAGATTCCTAGAGACCCGCTGTTTACAGCGGG 542
Db 1583 ACCGAACATGAGAAATACATACATCAGAGATTCCTAGAGACCCGCTGTTTACAGCGGG 1642
Qy 543 GTTTTCTTGTGACAAAGATCTCTCAATACCGCAGAGTCTAGACTGTGTGACTTC 602
Db 1643 GTTTTCTTGTGACAAAGATCTCTCAATACCGCAGAGTCTAGACTGTGTGACTTC 1702
Qy 603 TCTCAATTTTCTAGGGGAACTACCGTGTGTCTTGGCCAAATTCGCAAGTCCCAACTTC 662
Db 1703 TCTCAATTTTCTAGGGGAGTGCCTGTGTCTTGGCCCTAAATTCGCAAGTCCCAACTTC 1762
Qy 663 CAATCACTACCAACCTCTCTCTCTCCATCTGTCATGCTGCTGATGCTGATGCTGGG 722
Db 1763 CAATCACTACCAACCTCTCTCTCTCCATCTGTCATGCTGCTGATGCTGATGCTGGG 1822
Qy 723 GCCTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 782
Db 1823 GCCTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1882
Qy 783 GGAATCAAGATGATGTTGGCCGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 842
Db 1883 GGAATCAAGATGATGTTGGCCGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1942
Qy 843 GGAACCATGAGAGCCTGACGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 902
Db 1943 GGAACCATGAGAGCCTGACGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2002
Qy 903 CTGTACAAACCTTCCGATGGAACCTGATCTGTATCTCCATCCATCATCTCTGGCTTT 962
Db 2003 CTGTACAAACCTTCCGATGGAACCTGATCTGTATCTCCATCCATCATCTCTGGCTTT 2062
Qy 963 CGAAATTCCTATGAGAGTGGGCTCTGAGCCGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1022
Db 2063 CGAAATTCCTATGAGAGTGGGCTCTGAGCCGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2122
Qy 1023 ATTTTTCAGTGTGTTGAGGCTTCCGCCAGTCTTGGCTTCAATTATGATGAT 1082
Db 2123 ATTTTTCAGTGTGTTGAGGCTTCCGCCAGTCTTGGCTTCAATTATGATGAT 2182
Qy 1083 GTTGTACTGGGGCCCAAGTCTGTACACATCTTGAATGCCCTTTTACCGCTGTACCAAT 1142
Db 2183 GTTGTACTGGGGCCCAAGTCTGTACACATCTTGAATGCCCTTTTATACCGCTGTACCAAT 2242
Qy 1143 TTTCTTTTGTCTTTGGGTATACATTAAACCTTAATAA 1181
Db 2243 TTTCTTTTGTCTTTGGGTATACATTAAACCTTAATAA 2281

RESULT 9
US-09-718-095-8
; Sequence 8, Application US/09718095
; GENERAL INFORMATION:
; APPLICANT: Stuyver, Lieven
; APPLICANT: VAN GEYT, Sija
; APPLICANT: DE GENDT, Sija
; TITLE OF INVENTION: New HBV Sequences
; FILE REFERENCE: 2551-52
; CURRENT APPLICATION NUMBER: US/09/718.095
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: EP9870252.6
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US60/169,287
```

```

1  PRIOR FILING DATE: 1999-12-07
2  NUMBER OF SEQ ID NOS: 169
3  SOFTWARE: PatentIn Ver. 2.1
4  SEQ ID NO 8
5  LENGTH: 1203
6  TYPE: DNA
7  ORGANISM: Hepatitis B virus
8  FEATURE:
9  NAME/KEY: variation
10 LOCATION: (33)..(35)
11 OTHER INFORMATION: unknown or other
12 NAME/KEY: variation
13 LOCATION: (161)
14 OTHER INFORMATION: unknown or other
15
16 US-09-718-095-8

```

Query Match	85.5%	Score	1010.2	DB 31	Length	1203
Best Local Similarity	91.5%	Pred	No. 3.5e-286			
Matches 1069	Conservative	0	Mismatches	99	Indels	0
					Gaps	0

Qy	GGGGGGAATCTTTCCACAGCAATCTCTGGGATTTCTTCCGACACACAGTTGATTC	62
Db	36 GGGAAAGAACCTTTCCACGAGCAATCTCTAGATTTCTTCCGATCAGATTTGACC	95
Qy	63 AGCTTCAGACAAACAACAATCAGATTGGACTTCAATCCACAAGACACTG	122
Db	96 AGCATTCAGAGCAAAATACCAACAAATCAGATTGGAGCTTCAATCCAAAAGACCTTG	155
Qy	123 GCCAGACGCCAACAAGTATGAGCTGGAGCATTTGGACTGGGGTTACCCACCGCACGG	182
Db	156 GCCAGNAGCCAAACAGTATGAGTTGAGACCTATGACCCGGGTTACCCCTTCACACGG	215
Qy	183 AGGCTTTTGGGGTGVAGGCTTCAGGCTCAGGGCACTAAACAACAACCTTGACGACAAATCC	242
Db	216 AGGCTTTTGGGGTGVAGGCTTCAGTCTAGGGCACACTAACAACTTTGGCCAGAGATCC	275
Qy	243 GCTCTCTGCTCCACCAATGCGCAGTCAAGAAAGGACGCTACCCGCTGTCCACCTTT	302
Db	276 GCTCTCTGCTCCACCAATGCTCAGTCAAGGAGGACGCTACTCTCCACCACT	335
Qy	303 GAAACAACATCTATCTTCAGGCCATGAGTGAATCTCCAACTTTCCACAAATCTTGCA	362
Db	336 AAGAGACAGTCACTCTCAGGCCATGAGTGAATCTCAGAGATTCCACAAAGCTTACA	395
Qy	363 AGATCCAGAGTGAAGTCTGATTTCCCTGCTGATGAGCTCAGTTAGGAAAGTAAA	422
Db	396 AATATCCAAAGTAGGGGCTGTATTTCTCTGTGTGCTCAGTTAGGAAATGTGAA	455
Qy	423 CCTGTTCGCACTACTGTCTTCCATATGTCATCTTTCCAGAGATTGGGACCCCTGC	482
Db	456 CCTGTTCGCACTATTTGCTCTTCACTCTGTCATCTTTCCAGATTGGGACCCCTGC	515
Qy	483 GCGGAACATGAGAAACATCAATCAGGATTTCTAGAACCCCTGCTCGTGTTAACAGGGGG	542
Db	516 ACCGAACATGAGAAACATCAATCAGGATTTCTAGAACCCCTGCTCGTGTTAACAGGGGG	575
Qy	543 GTTTTCTTGTGACAAGAATCCTCAATACCGGAGTCTAGACTGTGATGAGACTTC	602
Db	576 GTTTTCTTGTGACAAGAATCCTCAATACCGGAGTCTAGACTGTGATGAGACTTC	635
Qy	603 TCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAAATTGCAAGTCCCACTC	662
Db	636 TCTCAATTTTCTAGGGGAGTGCCGTGTGTCTGGCCGTAATTGGCAGTCCCACTC	695
Qy	663 CAATCACTCAACCACTCTGTCTTCAACTTGTCTGTATATGCTGAGATGTCTGCG	722
Db	696 CAATCACTCAACCAATCTCTGTCTTCAACTTGTCTGTATATGCTGAGATGTCTGCG	755
Qy	723 GCGTTTATATCTTCTCTTCACTCTGCTGATAGCCATCTTGTGTTGTTCTT	782
Db	756 GCGTTTATATCTTCTCTTCACTCTGCTGATAGCCATCTTGTGTTGTTCTT	815
Qy	783 GGACTATCAAGTATGTGCGGTTTGTCTTAAATTCAGAGATTCAACACACAGAC	842

Db	816	GGACTTATCAGATATGTTGGCCGTTTGCTCTGATTCAGAGATCCTCGACCAACAAGTAC	875
Qy	843	GGGACCATGACAGACCCTGCACGACTCCGTGCTCAAGAACTCTATGATCCTCCTGTG	902
Db	876	GGGACCTCTGAAAACTGCACGACTCCTGTCTCAAGCAACTATATATCCTCATGTG	935
Qy	903	CTGTACAAAACCTTCGGATGGAACCTGACCTGTATTTCCGATCCGATCATCTGGGCTTT	962
Db	936	CTGTACAAAACCTTCGGACGGAAATTCGACCTGTATTTCCATCCGATCGACTGGGCTTT	995
Qy	963	CGAAAAATTCCTATGGAAGTGGGCTTAGGCCGTTTCTCCTGGCTCAGTTTACTAGTGC	1022
Db	996	CGAAAAATTCCTATGGAAGTGGGCTTAGGCCGTTTCTCCTGGCTCAGTTTACTAGTGC	1055
Qy	1023	ATTGTTCAGTGGTTCGTAGGGCTTTCCCCACAGCTTTGGCTTCAGTTATATGATGAT	1082
Db	1056	ATTGTTCAGTGGTTCGTAGGGCTTTCCCCACAGCTTCGGCTTTCAGCTATATGATGAT	1115
Qy	1083	GTTGTACTGGGGGCAAGCTCTGTACACACATCTGAATCCCTTTTACCGCTATACAAT	1142
Db	1116	GTTGTACTGGGGGCAAGCTCTGTACACACATCTGAATCCCTTTTATACCGCTATACAAT	1175
Qy	1143	TTTCTTTTGTCTTTGGGATACATTTAA	1170
Db	1176	TTTCTTTTGTCTTTGGGATACATTTAA	1203

```

RESULT 10
PCT-US02-32263-45
; Sequence 45, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Deprees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bowe, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO
; CURRENT APPLICATION NUMBER: PCT/US02/32263
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Hepatitis B virus
PCT-US02-32263-45

```

Query Match	85.1%;	Score 1005.2;	DB 1;	Length 1203;
Best Local Similarity	91.2%;	Pred. No. 1e-284;		
Matches 1067;	Conservative	0;	Mismatches 103;	Indels 0;
			Gaps	0





```

Qy 361 CAAGATCCAGAGTAGAGGTCTGATTTCCCTGCTGAGGCTCAGTTCAAGAAACAGTA 420
Db 394 CTAGATCCAGAGTAGAGGCGCTATATTTTCCGCTGGTGGCTCCAGTTCCGGAAACAGTA 453
Qy 421 AACCTGTTCCAGACTACTCTCTCTCCCATATGTCATTTTCGAGGATTGGGACCTT 480
Db 454 AACCTGTTCCAGACTACTCTCTCAACCATATGTCATTTTCGAGGATTGGGACCTT 513
Qy 481 GCGGGAACATGAGAAACATCAATCAGATTCTGAGACCCCTGCTGGTTACAGGGG 540
Db 514 GCACCGAACATGAGAAACATCAATCAGATTCTGAGACCCCTGCTGGTTACAGGGG 573
Qy 541 GGGTTTTCTGTTGACAGAAATCTCACAATAACGACAGTCTAGACTCTGAGTACT 600
Db 574 GGGTTTTCTGTTGACAGAAATCTCACAATAACGACAGTCTAGACTCTGAGTACT 633
Qy 601 TCTTCATATTTCTTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660
Db 634 TCTTCATATTTCTTAGGGGAGACCAACGTTCTGCGGCAAAATTCGAGTCCCAACC 693
Qy 661 TCCAACTACTACCAACCTCTGTCTGCAACTGTGCTGATTCGCTGAGTGTCTG 720
Db 694 TCCAACTACTACCAACCTCTGTCTGCAACTGTGCTGATTCGCTGAGTGTCTG 753
Qy 721 CGGCGTTTATCATCTTCTCTTCATCTGCTGCTAGTGCCTCATCTTCTGTTGTTCTT 780
Db 754 CGGCGTTTATCATATCTCTCTTCATCTGCTGCTAGTGCCTCATCTTCTGTTGTTCTT 813
Qy 781 CTGGAATCAAGATATGTTGCCCGTTGTCTCTAATTCAGATCTTCAACACACAGC 840
Db 814 CTGGAATCAAGATATGTTGCCCGTTGTCTCTAATTCAGATCTTCAACACACAGC 873
Qy 841 AGGGAGCAATGAGAGCGCTGAGCGAAGTCTGCTCAAGAAACCTCATATGATCCCTGCT 900
Db 874 AGGGAGCAATGAGAGCGCTGAGCGAAGTCTGCTCAAGAAACCTCATATGATCCCTGCT 933
Qy 901 TECTGTACAAAACCTTCGAGTGAAGAACTGCACTGTATTCCTATCCATCATCTGAGCT 960
Db 934 TECTGTACAAAACCTTCGAGTGAAGAACTGCACTGTATTCCTATCCATCATCTGAGCT 993
Qy 961 TTGGGAAATTCCTATGAGAGTGGGCTCAGCCCGTTTCTCTGCTCAGTTTACTAGTG 1020
Db 994 TTGGGAAATTCCTATGAGAGTGGGCTCAGCCCGTTTCTCTGCTCAGTTTACTAGTG 1053
Qy 1021 CCATTGTTCACTGTTGCTGAGGGCTTTCCCACTGTTGGCTTCACTATATGAGAG 1080
Db 1054 CCATTGTTCACTGTTGCTGAGGGCTTTCCCACTGTTGGCTTCACTATATGAGAG 1113
Qy 1081 ATGTGTACTGAGGAGCAAGTCTGTACACCATCTTGAGTCCCTTTTACCGCTGTTACCA 1140
Db 1114 ATGTGTACTGAGGAGCAAGTCTGTACACCATCTTGAGTCCCTTTTACCGCTGTTACCA 1173
Qy 1141 ATTTTCTTTTGTCTTGGGTATACATTAA 1170
Db 1174 ATTTTCTTTTGTCTTGGGTATACATTGA 1203

RESULT 12
US-10-410-897-45
; Sequence 45, Application US/10410897
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hayes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Gary
; TITLE OF INVENTION: FACTOR IX: REMODELING AND GLYCOSYLATION OF FACTOR IX
; FILE REFERENCE: 040853-01-5058
; CURRENT APPLICATION NUMBER: US/10/410,897
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523

```

```

; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Hepatitis B virus
US-10-410-897-45

Query Match      85.1%; Score 1005.2; DB 52; Length 1203;
Best Local Similarity 91.2%; Pred. No. 1e-284;
Matches 1067; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 1 ATGGGGGCAAACTTTCCACCAAGCAATCCTCGGGGATCTTCCGACCAAGTTGAT 60
Db 34 ATGGGGAAGAACTTTCTGTTCTCCAAATCTCTGGGATTTCTTCCGATACCAAGTTGAC 93
Qy 61 CCAGCTTCAGAGCAAAACCAACCAATCCAGATTGGGACTTCAATCCCAACAGACACC 120
Db 94 CCGGCTTCAGAGCAAAACCAACCAATCCAGATTGGGACTTCAATCCCAACAGATCAC 153
Qy 121 TGGCCAGAGCCCAACAGATGAGCTGAGAGCTTGGACTGGGGTTACACCCACCGAC 180
Db 154 TGGCCAGAGCAATCAAGATGAGAGCGGAGACTTCGGGCCAGGGTTACACCCACACAC 213
Qy 181 GGAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGCATACACAAACCTTGCACCAAT 240
Db 214 GGCGGTCTTTGGGGTGAAGCCCTCAGGCTCAGGGCATATGACAAAGTGCACAGCG 273
Qy 241 CCGCTCTGCTTCCACCAATGCGCAGTCAAGAAAGCAGCTACCCGCTGCTCCACT 300
Db 274 CCGCTCTGCTTCCACCAATGCGCAGTCAAGAAAGCAGCTACCCGCTGCTCCACT 333
Qy 301 TTGAGAAACATCATCTCTCAAGCCATGAGTGAAGCTCCACAACTTTCACCAACCTG 360
Db 334 CTAGAGAGACATCATCTCTCAAGCCATGAGTGAAGCTCCACAACTTTCACCAACCTG 393
Qy 361 CAAGATCCAGAGTAGAGGTCTGATTTCCCTGCTGAGGCTCAGTTCAAGAAACAGTA 420
Db 394 CTAGATCCAGAGTAGAGGCGCTATATTTTCCGCTGGTGGCTCCAGTTCCGGAAACAGTA 453
Qy 421 AACCTGTTCCAGACTACTGTCTCTCCCATATGTCATTTTCGAGGATTGGGACCTT 480
Db 454 AACCTGTTCCAGACTACTGTCTCTCAACCATATGTCATTTTCGAGGATTGGGACCTT 513
Qy 481 GCGGGAACATGAGAAACATCAATCAGATTCTGAGACCCCTGCTGGTTACAGGGG 540
Db 514 GCACCGAACATGAGAAACATCAATCAGATTCTGAGACCCCTGCTGGTTACAGGGG 573
Qy 541 GGGTTTTCTGTTGACAGAAATCTCACAATAACGACAGTCTAGACTCTGAGTACT 600
Db 574 GGGTTTTCTGTTGACAGAAATCTCACAATAACGACAGTCTAGACTCTGAGTACT 633
Qy 601 TCTTCATATTTCTTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660
Db 634 TCTTCATATTTCTTAGGGGAGACCAACGTTCTGCGGCAAAATTCGAGTCCCAACC 693
Qy 661 TCCAACTACTACCAACCTCTGTCTGCAACTGTGCTGATTCGCTGAGTGTCTG 720
Db 694 TCCAACTACTACCAACCTCTGTCTGCAACTGTGCTGATTCGCTGAGTGTCTG 753

```

QY 721 CGCGGTTTATCATCTTCTCTCATCTCTGCTATGCTCATCTTCTTGTTGTTCTT 780  
| | | | |  
Db 754 CGCGGTTTATCATCTTCTCTCATCTCTGCTATGCTCATCTTCTTGTTGTTCTT 813  
| | | | |  
QY 781 CTGACATATCAAGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAAGC 840  
| | | | |  
Db 814 CTGACATCAACAGTATGTTGCCGTTTGTCTCTAATTCAGAGATCAATCAACCAAGC 873  
| | | | |  
QY 841 ACGGACATGACAGAGCTCTGCTCAAGAACTCTATGTAATCCCTCTGT 900  
| | | | |  
Db 874 ACGGACATGACAGAGCTCTGCTCAAGAACTCTATGTAATCCCTCTGT 933  
| | | | |  
QY 901 TGTGTACAAACCTTGATGAGAACTGACCTGTAATCCATCCATCATCTGGCT 960  
| | | | |  
Db 934 TGTGTACAAACCTTGATGAGAACTGACCTGTAATCCATCCATCATCTGGCT 993  
| | | | |  
QY 961 TTGGGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGAGTCACTTAACTAGT 1020  
| | | | |  
Db 994 TTGGCAAGATTTCTATGAGAGTGGGCTCAGTCCGTTTCTCTGAGTCACTTAACTAGT 1053  
| | | | |  
QY 1021 CCAATTGTTAGTGTCTGATGAGGCTTCCGCCCACTGTTGGCTTTCAGTTATATGATG 1080  
| | | | |  
Db 1054 CCAATTGTTAGTGTCTGATGAGGCTTCCGCCCACTGTTGGCTTTCAGTTATATGATG 1113  
| | | | |  
QY 1081 ATGTTGATGAGGAGGCAAGTCTGTACACCATCTTGAGTCCCTTTTACCCTGTTACCA 1140  
| | | | |  
Db 1114 ATGTTGATGAGGAGGCAAGTCTGTACACCATCTTGAGTCCCTTTTACCCTGTTACCA 1173  
| | | | |  
QY 1141 ATTTTCTTTTGTCTTGGGTATACATTAA 1170  
| | | | |  
Db 1174 ATTTTCTTTTGTCTTGGGTATACATTGA 1203  
| | | | |

## RESULT 13

US-10-410-913-45  
; Sequence 45; Application US/10410913  
; GENERAL INFORMATION:  
; APPLICANT: Neose Technologies, Inc.  
; APPLICANT: Defrees, Shawn  
; APPLICANT: Zopf, David  
; APPLICANT: Bayer, Robert  
; APPLICANT: Hakes, David  
; APPLICANT: Chen, Xi  
; APPLICANT: Howe, Caryn  
; TITLE OF INVENTION: GLYCOCONJUGATION METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE  
; FILE REFERENCE: 040853-01-5081  
; CURRENT APPLICATION NUMBER: US/10/410,913  
; PRIOR FILING DATE: 2003-04-09  
; PRIOR APPLICATION NUMBER: US 60/328,523  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/344,692  
; PRIOR FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: US 60/387,292  
; PRIOR FILING DATE: 2002-06-07  
; PRIOR APPLICATION NUMBER: US 60/391,777  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/396,594  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: US 60/404,249  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: US 60/407,527  
; PRIOR FILING DATE: 2002-08-28  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 45  
; LENGTH: 1203  
; TYPE: DNA  
; ORGANISM: Hepaticis B virus  
US-10-410-913-45

Query Match 85.1%; Score 1005.2; DB 52; Length 1203;  
Best Local Similarity 91.2%; Pred. No. 1e-284;

	Matches 1067;	Conservative 0;	Mismatches 103;	Indels 0;	Gaps 0;
QY	1	ATGGGACGAATCTTTCACACGAACTCTGAGATTTCTTCCGACCAAGTGTGAT	60		
Db	34	ATGGGACGAATCTTTCGATCCCAATCTCTGGGATTTCTTCCGATCAACAGTGGAG	93		
QY	61	CCAGCTTCAGAGCAACCAACCAATCAAGATTTGGACTTCATTCACCAAGACACC	120		
Db	94	CTGGCTTCGAGCCCAACTCAACCAATTCAGATTTGGACTTCACCAAGATTCAC	153		
QY	121	TGGCCAGAGCCCAACCAAGATTCAGATTTGGACTTCAGATTTGGACTTCAGATTT	180		
Db	154	TGGCCAGAGCCCAACCAAGATTCAGATTTGGACTTCAGATTTGGACTTCAGATTT	213		
QY	181	GGAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGAGTAAACAACTTGGCCAGCAAT	240		
Db	214	GGGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGAGTAAATGACAAAGTCCAGAGAG	273		
QY	241	CGGCTCTGCTTCCACCAATGCGCAGTACAGAGGCAAGCTTACCCGCTGTCTCACT	300		
Db	274	CTCTCTCTGTTTCCACCAATGCGCAGTACAGAGGCAAGCTTACCCGCTGTCTCACT	333		
QY	301	TTGAGAAACATCATCTCTCAAGGATGAGTGAATCCACCACTTTCACCAACATCTG	360		
Db	334	TTGAGAAACATCATCTCTCAAGGATGAGTGAATCCACCACTTTCACCAACATCTG	393		
QY	361	CAAGATCCCAAGTGAAGGCTGTATTTCCCTGCTGAGTCCAGTTCCAGAACAGTA	420		
Db	394	CTAGATCCCAAGTGAAGGCTGTATTTCCCTGCTGAGTCCAGTTCCAGAACAGTA	453		
QY	421	AACCTGTTCCGACTACTGTCTCTCCATATCTCAATCTTCTCGAGATTTGGGACCT	480		
Db	454	AACCTGTTCCGACTACTGTCTCTCCATATCTCAATCTTCTCGAGATTTGGGACCT	513		
QY	481	GGCGGACATGAGAGAAATCAATCAAGATTCCTGAGACCCCTGCTCGTTACAGGGG	540		
Db	514	GCACCAACATGAGAGAAATCAATCAAGATTCCTGAGACCCCTGCTCGTTACAGGGG	573		
QY	541	GGGTTTTCTTGTGACCAAGATCTCAACATTCAGGAGTCTGAGTCTGAGTCTGAGT	600		
Db	574	GGGTTTTCTTGTGACCAAGATCTCAACATTCAGGAGTCTGAGTCTGAGTCTGAGT	633		
QY	601	TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC	660		
Db	634	TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC	693		
QY	661	TCCAAATCACTACCAACCTCTCTCTCAACCTTCTGCTGATATGCTGAGATGCTG	720		
Db	694	TCCAAATCACTACCAACCTCTCTCTCAACCTTCTGCTGATATGCTGAGATGCTG	753		
QY	721	CGCGGTTTATCATCTTCTCTCATCTGCTCTAATTCAGAGATCTTCAACCAAGC	780		
Db	754	CGCGGTTTATCATCTTCTCTCATCTGCTCTAATTCAGAGATCTTCAACCAAGC	813		
QY	781	CTGACATCAAGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAAGC	840		
Db	814	CTGACATCAAGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAAGC	873		
QY	841	ACGGACATGACAGAGCTCTGCTCAAGAACTCTATGTAATCCCTCTGT	900		
Db	874	ACGGACATGACAGAGCTCTGCTCAAGAACTCTATGTAATCCCTCTGT	933		
QY	901	TGTGTACAAACCTTGATGAGAACTGACCTGTAATCCATCCATCATCTGGCT	960		
Db	934	TGTGTACAAACCTTGATGAGAACTGACCTGTAATCCATCCATCATCTGGCT	993		
QY	961	TTGGGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGAGTCACTTAACTAGT	1020		
Db	994	TTGGCAAGATTTCTATGAGAGTGGGCTCAGTCCGTTTCTCTGAGTCACTTAACTAGT	1053		
QY	1021	CCAATTGTTAGTGTCTGATGAGGCTTCCGCCCACTGTTGGCTTTCAGTTATATGATG	1080		
Db	1054	CCAATTGTTAGTGTCTGATGAGGCTTCCGCCCACTGTTGGCTTTCAGTTATATGATG	1113		

QY 1081 ATGTGTACTGGGGGCAAGTCTGTACACCATCTTGTAGTCCCTTTTACCGCTTTACCA 1140  
| | | | |  
DB 1114 ATGTGTATGGGGGCAAGTCTGTACACCATCTTGTAGTCCCTTTTACCGCTTTACCA 1173  
| | | | |  
QY 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170  
| | | | |  
DB 1174 ATTTCTTTTGTCTTTGGGTATACATTGA 1203  
| | | | |  
  
RESULT 14  
US-10-410-930-45  
; Sequence 45, Application US/10410930  
; GENERAL INFORMATION:  
; APPLICANT: Neose Technologies, Inc.  
; APPLICANT: Defrees, Shawn  
; APPLICANT: Zopf, David  
; APPLICANT: Bayer, Robert  
; APPLICANT: Hakes, David  
; APPLICANT: Chen, Xi  
; APPLICANT: Bove, Caryn  
; TITLE OF INVENTION: INTERFERON BETA: REMODELING AND GLYCOCONJUGATION OF INTERFERON  
; TITLE OF INVENTION: BETA  
; FILE REFERENCE: 040853-01-5056  
; CURRENT APPLICATION NUMBER: US/10/410,930  
; PRIOR FILING DATE: 2003-04-09  
; PRIOR APPLICATION NUMBER: US 60/328,523  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/344,692  
; PRIOR FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: US 60/387,292  
; PRIOR FILING DATE: 2002-06-07  
; PRIOR APPLICATION NUMBER: US 60/391,777  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/396,594  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: US 60/404,249  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: US 60/407,527  
; PRIOR FILING DATE: 2002-08-28  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 45  
; LENGTH: 1203  
; TYPE: DNA  
; ORGANISM: Hepatitis B virus  
US-10-410-930-45  
  
Query Match 85.1%; Score 1005.2; DB 52; Length 1203;  
Best Local Similarity 91.2%; Pred. No. 1e-284;  
Matches 1067; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

DB 334 CTAAAGAGACATCTCAGGCCATGAGTGAAGTCCACAACAATTCACCAAGCTCTG 393  
| | | | |  
QY 361 CAAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGTGTGCTCCAGTTACAGAAACAGTA 420  
| | | | |  
DB 394 CTAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGTGTGCTCCAGTTACAGAAACAGTA 453  
| | | | |  
QY 421 AACCGTGTCCGACTACTGTCTCTCCATATCGTCAATCTTCTCGAGATTTGGGAGCCCT 480  
| | | | |  
DB 454 AACCGTGTCCGACTACTGTCTCTCCATATCGTCAATCTTCTCGAGATTTGGGAGCCCT 513  
| | | | |  
QY 481 GCGCGAATCATGAGAACATCAATCAGATTTCTAGAACCCCTGCTCTGTATTAAGGGG 540  
| | | | |  
DB 514 GCACCGAATCATGAGAACCAATCAGATTTCTAGAACCCCTGCTCTGTATTAAGGGG 573  
| | | | |  
QY 541 GGGTTTTTCTTTGTGAAGAATCTCTCAATATCCGACAGTCTGACTCTGTGTGACT 600  
| | | | |  
DB 574 GGGTTTTTCTTTGTGAAGAATCTCTCAATATCCGACAGTCTGACTCTGTGTGACT 633  
| | | | |  
QY 601 TCTCTCAATTTCTAGGGGGAACTACCGTGTCTTGGCAAAATTCGACATCCCAACC 660  
| | | | |  
DB 634 TCTCTCAATTTCTAGGGGGAACTACCGTGTCTTGGCAAAATTCGACATCCCAACC 693  
| | | | |  
QY 661 TCCAATCACTCAACAACCTCTCTCTCAACTTGTCTGTATTCGTGATGTGTCTG 720  
| | | | |  
DB 694 TCCAATCACTCAACAACCTCTCTCTCAACTTGTCTGTATTCGTGATGTGTCTG 753  
| | | | |  
QY 721 CGCGTTTTATATCTTCTCTTCAATCTCTGTCTATGCTCTATCTTCTTGTGTTCTT 780  
| | | | |  
DB 754 CGCGTTTTATATCTTCTCTTCAATCTCTGTCTATGCTCTATCTTCTTGTGTTCTT 813  
| | | | |  
QY 781 CTGSACTATCAAGTATGTGCGCGTGTCTCTCAATTTCAAGATCTTCAACCAACCAC 840  
| | | | |  
DB 814 CTGSACTATCAAGTATGTGCGCGTGTCTCTCAATTTCAAGATCTTCAACCAACCAC 873  
| | | | |  
QY 841 ACGGACCATGAGAGCTCTGACGATCTCTGCTCAAGAACCTCTATGATCTCTCTGT 900  
| | | | |  
DB 874 ACGGACCATGAGAGCTCTGACGATCTCTGCTCAAGAACCTCTATGATCTCTCTGT 933  
| | | | |  
QY 901 TGTCTTACAAAACCTTCTGATGAGAACTGACCTGTATTTCCATCCATCTCTGGCT 960  
| | | | |  
DB 934 TGTCTTACAAAACCTTCTGATGAGAACTGACCTGTATTTCCATCCATCTCTGGCT 993  
| | | | |  
QY 961 TTGGAAATTCCTATGAGAGTGGGCTCAGCGCTTCTCTGTGCTAGTTACTAGTG 1020  
| | | | |  
DB 994 TTGGAAATTCCTATGAGAGTGGGCTCAGCGCTTCTCTGTGCTAGTTACTAGTG 1053  
| | | | |  
QY 1021 CCATTGTTCAGTGTCTGAGGGCTTCCCACTGTTTGGCTTCAATTATGATG 1080  
| | | | |  
DB 1054 CCATTGTTCAGTGTCTGAGGGCTTCCCACTGTTTGGCTTCAATTATGATG 1113  
| | | | |  
QY 1081 ATGTGTACTGGGGGCAAGTCTGTACACCATCTTGTAGTCCCTTTTACCGCTTTACCA 1140  
| | | | |  
DB 1114 ATGTGTATGGGGGCAAGTCTGTACACCATCTTGTAGTCCCTTTTACCGCTTTACCA 1173  
| | | | |  
QY 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170  
| | | | |  
DB 1174 ATTTCTTTTGTCTTTGGGTATACATTGA 1203  
| | | | |  
  
RESULT 15  
US-10-410-945-45  
; Sequence 45, Application US/10410945  
; GENERAL INFORMATION:  
; APPLICANT: Neose Technologies, Inc.  
; APPLICANT: Defrees, Shawn  
; APPLICANT: Zopf, David  
; APPLICANT: Bayer, Robert  
; APPLICANT: Hakes, David  
; APPLICANT: Chen, Xi  
; APPLICANT: Bove, Caryn  
; TITLE OF INVENTION: ERYTHROPOIETIN: REMODELING AND GLYCOCONJUGATION OF ERYTHROPOIETIN  
; FILE REFERENCE: 040853-01-5083

```

; CURRENT APPLICATION NUMBER: US/10/410,945
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Hepatitis B virus
US-10-410-945-45

Query Match      85.1%; Score 1005.2; DB 52; Length 1203;
Best Local Similarity 91.2%; Pred. No. 1e-284;
Matches 1067; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY      1  ATGGGCGAGATCTTTCCACAGCAATCTCTGGGATTTCTTCCGACCAACGATTGGAT 60
DB      34  ATGGGCGAGATCTTTCTGTCTTCCTCAATCTCTGGGATTTCTTCCGATCCAGTGGAC 93
QY      61  CCAGCCTTCAGAGCAACCAACCAATCCAGATTGGACTTCAATCCCAACAGACACC 120
DB      94  CTGCGCTTGGAGCCCACTCAACCAATCCAGATTGGACTTCAATCCCAACAGATCAC 153
QY      121  TGGCCAGAGCCCAACCAAGTAGAGCTGAGCAATTCGAGCTGGGGTTACCCCAACCGCAC 180
DB      154  TGGCCAGAGGCAATCAAGGTAGAGCGGAGACTTTCGGGCAAGGTTCAACCCCAACAC 213
QY      181  GGAGGCTTTTGGGTTGAGAGCCCTCAGGCTCAGGGGCTTAACAAACCTTCCAGCAAT 240
DB      214  GGGGCTTTTGGGTTGAGAGCCCTCAGGCTCAGGGGCTTAACAAACCTTCCAGCAAGCG 273
QY      241  CCGCCTCTCTCTTCCACCAATCGCAGTCAAGAGGAGGAGGCTACCCCGCTGTCTCACT 300
DB      274  CTTCTCTCTCTTCCACCAATCGGAGTCAAGAGGAGGAGGCTACCTCCATCTCTCACT 333
QY      301  TTGAGAAACATCATCTCAAGCCATGAGTGAATCTCAACACTTTCACCAAACTCTG 360
DB      334  CTAAAGAGACAGTATCTCAGGCAATGAGTGAATCTCAACAACTTTCACCAAGCTCTG 393
QY      361  CAAGATCCAGAGTGAAGGTCTGTATTTCCCTGCTGTGGCTTCAATTCAGGAACGTA 420
DB      394  CTAGATCCAGAGTGAAGGCTGTATTTTCTGCTGTGGCTTCAATTCAGGAACGTA 453
QY      421  AACCTTTCAGACTCTCTCTCCATATCGTCAATCTTCTCGAGATTGGGAGCCCT 480
DB      454  AACCTTTCAGACTCTCTCTCCATATCGTCAATCTTCTCGAGACTGGGAGCCCT 513
QY      481  GCGGGAACATGAGAACATCATCAGATTCCTAGAGCCCTGCTGTGTTCAGAGCG 540
DB      514  GCACGGAACATGAGAGCAACATCAGGATTCCTAGAGCCCTGCTGTGTTCAGAGCG 573
QY      541  GGGTTTTTCTTGTGACAAAGATCTCAATATCCGAGAGTCTAGACTGTGTGACT 600
DB      574  GGGTTTTTCTTGTGACAAAGATCTCAATATCCAGAGTCTAGACTGTGTGACT 633
QY      601  TCTCTCAATTTTCTAGGGGGAACATACCTGTGTCTTGGCCAAATTCGAGTCCCAACC 660
DB      634  TCTCTCAATTTTCTAGGGGGAACATACCTGTGTCTTGGCCAAATTCGAGTCCCAACC 693
QY      661  TCGAATCATCTACCAACTCTGTCTCTCAACTTGTGTCTGTATTCGCTGGAATGTCTG 720
```

```

DB      694  TCGAATCATCTACCAACTCTGTCTCTCAACTTGTGTCTGTATTCGCTGGAATGTCTG 753
QY      721  CGGCTTTATATATCTTCCCTTCAATCTGCTGATAGCTCATCTTCTGTGTTCTT 780
DB      754  GGGCTTTATATATCTTCCCTTCAATCTGCTGATAGCTCATCTTCTGTGTTCTT 813
QY      781  CTGACATATCAAGATATGTTGCCGTTTCTCTATATTCAGAGATCTTCAACACACAG 840
DB      814  CTGACATATCAAGATATGTTGCCGTTTCTCTATATTCAGAGATCTTCAACACACAG 873
QY      841  ACCGGAACATGCAAGACTCTGACAGATCTCTGCTCAAGAACTCTATATCTCTCTG 900
DB      874  ACCGGAACATGCAAGACTCTGACAGATCTCTGCTCAAGAACTCTATATCTCTCTG 933
QY      901  TGCTGTACAAACCTTCCGATGGAACCTGACACCTGTATTCATCCATCATCTGGGCT 960
DB      934  TGCTGTACAAACCTTCCGATGGAACCTGACACCTGTATTCATCCATCATCTGGGCT 993
QY      961  TTGGAATAATCTATGAGAGTGGGCTGAGCCGTTTCTCTGCTCAGTTACTAGTG 1020
DB      994  TTGGAATAATCTATGAGAGTGGGCTGAGCCGTTTCTCTGCTCAGTTACTAGTG 1053
QY      1021  CCATTGTTCAAGTGTGTTCTGAGGCTTTCCCACTGTTTGGCTTCAATATAGAGTG 1080
DB      1054  CCATTGTTCAAGTGTGTTCTGAGGCTTTCCCACTGTTTGGCTTCAATATAGAGTG 1113
QY      1081  ATGTTGACTGGGGGCAAGTCTGTACACATCTTGAATCCCTTTTACCGCTGTACCA 1140
DB      1114  ATGTTGACTGGGGGCAAGTCTGTACACATCTTGAATCCCTTTTACCGCTGTACCA 1173
QY      1141  ATTTCTTTTCTTTGGTATATATTTAA 1170
DB      1174  ATTTCTTTTCTTTGGTATATATTTGA 1203
```

Search completed: September 15, 2003, 03:47:00  
Job time : 2814 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 15, 2003, 01:33:11; Search time 272 Seconds  
(without alignments)  
3647.246 Million cell updates/sec

Title: US-09-821-877-1  
Perfect score: 1181  
Sequence: 1 atgsggcagaaattccac.....tacatttaaccctaataa 1181

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 604708 seqs, 420004563 residues  
Total number of hits satisfying chosen parameters: 1209416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_NA\_New:\*  
1: /cgn2\_6/pdata/1/pna/PCT\_NEW\_COMB.seq:\*  
2: /cgn2\_6/pdata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/pdata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/pdata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/pdata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/pdata/1/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/pdata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	810.8	68.7	846	6	US-10-383-317-11 Sequence 11, Appl
2	756.4	64.0	846	6	US-10-383-317-9 Sequence 9, Appl
3	710.8	60.2	1977	1	PCT-US03-18993-97 Sequence 97, Appl
4	412	34.9	849	6	US-10-383-317-15 Sequence 15, Appl
5	124.8	10.6	128	1	PCT-US03-18993-1 Sequence 1, Appl
6	88.2	7.5	93	1	PCT-US03-18993-68 Sequence 68, Appl
7	71.4	6.0	73	1	PCT-US03-18993-69 Sequence 69, Appl
8	69.4	5.9	71	1	PCT-US03-18993-67 Sequence 67, Appl
9	57.4	4.9	59	1	PCT-US03-18993-95 Sequence 95, Appl
10	53.8	4.6	57	6	US-10-388-918A-7 Sequence 7, Appl
11	50.8	4.3	54	1	PCT-US03-18993-4 Sequence 4, Appl
12	50	4.2	50	1	PCT-US03-18993-84 Sequence 84, Appl
13	48.4	4.1	50	1	PCT-US03-18993-2 Sequence 2, Appl
14	43	3.6	45	1	PCT-US03-18993-78 Sequence 78, Appl
15	42.8	3.6	46	1	PCT-US03-18993-142 Sequence 142, App
16	42	3.6	42	1	PCT-US03-18993-83 Sequence 83, Appl
17	40	3.4	40	1	PCT-US03-18993-74 Sequence 74, Appl
18	40	3.4	40	1	PCT-US03-18993-77 Sequence 77, Appl
19	38	3.2	23261	7	US-60-495-114-16509 Sequence 16509, A
20	38	3.2	96592	1	PCT-US02-36071A-175 Sequence 175, App
21	37.6	3.2	58398	7	US-60-495-135-3692 Sequence 3692, Ap
22	37.4	3.2	39	1	PCT-US03-18993-90 Sequence 90, Appl
23	36.8	3.1	201	7	US-60-495-114-38006 Sequence 38006, A
24	36.8	3.1	201	7	US-60-495-114-38007 Sequence 38007, A
25	36.8	3.1	201	7	US-60-495-114-38009 Sequence 38009, A
26	36.8	3.1	16701	7	US-60-487-610-19270 Sequence 19270, A

C 27	36.8	3.1	57978	7	US-60-487-610-19974	Sequence 19974, A
C 28	36.8	3.1	384485	7	US-60-495-114-16382	Sequence 16382, A
C 29	36.2	3.1	3379	7	US-60-487-610-1316	Sequence 1316, Ap
C 30	36.2	3.1	3379	7	US-60-485-450-843	Sequence 843, App
C 31	36	3.0	36	1	PCT-US03-18993-70	Sequence 70, Appl
C 32	36	3.0	36	1	PCT-US03-18993-82	Sequence 82, Appl
C 33	36	3.0	1072	7	US-60-493-007-3096	Sequence 3096, Ap
C 34	35.4	3.0	27587	1	PCT-US02-38582-253	Sequence 253, App
C 35	35.4	3.0	52846	7	US-60-487-610-19654	Sequence 19654, A
C 36	35.4	3.0	13831263	5	US-09-947-914-41	Sequence 41, Appl
C 37	35.2	3.0	2037	6	US-10-425-114A-26697	Sequence 26697, A
C 38	35.2	3.0	13508	7	US-60-487-610-19923	Sequence 19923, A
C 39	35.2	3.0	13831263	5	US-09-947-914-41	Sequence 41, Appl
C 40	35	3.0	35	1	PCT-US03-18993-89	Sequence 89, Appl
C 41	35	3.0	35	1	PCT-US03-18993-92	Sequence 92, Appl
C 42	35	3.0	35	1	PCT-US03-18993-126	Sequence 126, App
C 43	35	3.0	30400	7	US-60-487-610-19473	Sequence 19473, A
C 44	34.8	2.9	923	6	US-10-425-114A-30410	Sequence 30410, A
C 45	34.8	2.9	927	6	US-10-425-114A-31990	Sequence 31990, A

ALIGNMENTS

RESULT 1  
US-10-383-317-11  
Sequence 11, Application US/10383317  
GENERAL INFORMATION:  
APPLICANT: Punnonen, Juha  
APPLICANT: Baas, Steven H.  
APPLICANT: Whalen, Robert Gerald  
APPLICANT: Howard, Russell  
APPLICANT: Stemmer, Willem P.C.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Antigen library Immunization  
FILE REFERENCE: 018097-028710US  
CURRENT APPLICATION NUMBER: US/10/383,317  
CURRENT FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: US/09/247,890  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/074,294  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-11  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/105,509  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-23  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 11  
LENGTH: 846  
TYPE: DNA  
ORGANISM: Hepatitis B virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(846)  
OTHER INFORMATION: Pres2-S coding region of hepatitis B virus ayw  
OTHER INFORMATION: surface antigen (HBsAg)  
US-10-383-317-11  
Query Match 68.7%; Score 810.8; DB 6; Length 846;  
Best Local Similarity 97.4%; Pred. No. 3.4e-254;  
Matches 824; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 325 ATGCGATGGAATTCACCAACTTTCACCAAACTGCAAGATCCAGAGTGAGAGTCG 384  
1 ATGCGATGGAATTCACCAACTTTCACCAAACTGCAAGATCCAGAGTGAGAGTCG 60  
QY 385 TATTTCCCTGCTGCTGCTCCAGTTCCAGAAACAGTAACCCCTGTTCCGACTACTGCTCT 444  
61 TATTTCCCTGCTGCTGCTCCAGTTCCAGAAACAGTAACCCCTGTTCCGACTACTGCTCT 120  
QY 445 CCCATATGTCATATTTCTTCAGAGATTGGGACCCCTGGCGGAGATGAGAACTACCA 504  
121 CCTTATGTCATATTTCTTCAGAGATTGGGACCCCTGGCGGAGATGAGAACTACCA 180

```
QY 505 TCAGATTCCTAGGACCCCTGCTGCTGTTACAGCGGGGTTTTCTTGTGACAGATC 564
DB 181 TCAGATTCCTAGGACCCCTGCTGCTGTTACAGCGGGGTTTTCTTGTGACAGATC 240
QY 565 CTCACATACCGGAGAGCTAGACTCGTGGGAGCTCTCTCAATTTTCTAGGGGAACT 624
DB 241 CTCACATACCGGAGAGCTAGACTCGTGGGAGCTCTCTCAATTTTCTAGGGGAACT 300
QY 625 ACCGTGTGCTTGGCCAAAATTCGAGTCCCAACCTCAATCACTGACCACTCTCTGT 684
DB 301 ACCGTGTGCTTGGCCAAAATTCGAGTCCCAACCTCAATCACTGACCACTCTCTGT 360
QY 685 CTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
DB 361 CTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 745 ATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
DB 421 ATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 805 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
DB 481 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 865 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
DB 541 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 925 AACTGCACTGTAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
DB 601 AACTGCACTGTAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 985 GCCTGACCCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
DB 661 GCCTGACCCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1045 CTTTCCCCCACTGTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGATG 1104
DB 721 CTTTCCCCCACTGTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGATG 780
QY 1105 TACACATCTTGAGTCCCTTTTTCAGCTGTTACCAATTTTCTTGTGCTTGGGTATAC 1164
DB 781 TACACATCTTGAGTCCCTTTTTCAGCTGTTACCAATTTTCTTGTGCTTGGGTATAC 840
QY 1165 ATTTAA 1170
DB 841 ATTTAA 846
RESULT 2
US-10-383-317-9
; Sequence 9, Application US/10383317
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha
; APPLICANT: Baas, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/10/383,317
; PRIOR APPLICATION NUMBER: 2003-03-07
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/074,294
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-11
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/105,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 846
```

```
; TYPE: DNA
; ORGANISM: Hepatitis B virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(846)
; OTHER INFORMATION: PreS2-S coding region of hepatitis B virus adr
; OTHER INFORMATION: surface antigen (HBsAg)
US-10-383-317-9
Query Match 64.0%; Score 756.4; DB 6; Length 846;
Best Local Similarity 93.4%; Pred. No. 2e-236;
Matches 790; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 325 ATGCAAGTGAATCTCAACAACTTTCCACCAATCTGCAAGATCCGAGTGAAGTCTG 384
DB 1 ATGCAAGTGAATCTCAACAACTTTCCACCAATCTGCAAGATCCGAGTGAAGTCTG 60
QY 385 TATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
DB 61 TATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 445 CCATATCTGCAATCTTCTGAGAGATGGGAGCCCTGCGCGAATGAGAAATCATACA 504
DB 121 CCATATCTGCAATCTTCTGAGAGATGGGAGCCCTGCGCGAATGAGAAATCATACA 180
QY 505 TCAGATTCCTAGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
DB 181 TCAGATTCCTAGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 565 CTCACAAATCCGAGAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624
DB 241 CTCACAAATCCGAGAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 625 ACCGTGTGCTTGGCCAAAATTCGAGTCCCAACCTCAATCACTGACCACTCTCTGT 684
DB 301 ACCGTGTGCTTGGCCAAAATTCGAGTCCCAACCTCAATCACTGACCACTCTCTGT 360
QY 685 CTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
DB 361 CTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 745 ATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
DB 421 ATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 805 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
DB 481 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 865 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
DB 541 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 925 AACTGCACTGTAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
DB 601 AACTGCACTGTAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 985 GCCTGACCCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
DB 661 GCCTGACCCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1045 CTTTCCCCCACTGTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGATG 1104
DB 721 CTTTCCCCCACTGTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGATG 780
QY 1105 TACACATCTTGAGTCCCTTTTTCAGCTGTTACCAATTTTCTTGTGCTTGGGTATAC 1164
DB 781 TACACATCTTGAGTCCCTTTTTCAGCTGTTACCAATTTTCTTGTGCTTGGGTATAC 840
QY 1165 ATTTAA 1170
DB 841 ATTTAA 846
```

```
RESULT 3
PCT-US03-18993-97
; Sequence 97, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Himen, Jeffery M.
; APPLICANT: Kolk, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Germain, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loy, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GPI34-02 UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 97
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-97
```

```
Query Match      60.2%; Score 710.8; DB 1; Length 1977;
Best Local Similarity 63.9%; Pred. No. 2.4e-221;
Matches 523; Conservative 228; Mismatches 67; Indels 0; Gaps 0;
```

```
QY 364 GATCCAGAGTGAAGGTGCTATTTCCTGCTGGTGGCTCCAGTTACAGAAACAGTAAAC 423
DB 17 GAUCCAGAGUAGGGGUGUUAUUCUUGGUGGUGUCCAGUUCAGAAACAGAAAC 76
QY 424 CCTGTTCCGACTACTGCTCTCTCCATATCGTCAATCTTCTCAGAGATTGGGAGCCCTGCG 483
DB 77 CCUGCUGCGAAUUGGCUUCUACAUUCUGUACAUUCUGCGAGAGCUGGGAGCCUUG 136
QY 484 CGGAACATGAGAAACATCATCAGAGATTCTTGAACCCCTGCTGTTACAGGCGGG 543
DB 137 ACGAUCAGGAAACAUCAUCAGAGAUUCCUAGAGCCCUUCUUCUGUACAGGCGGG 196
QY 544 TTTTCTGTTGACAAGATCCCAATACCGGAGAGTCTAGCTCGGAGTCTTCT 603
DB 197 UUUUUUUUUUAGCAAGAUUCCUACAACGCAAGUCUAGUUGGUGAGUUCU 256
QY 604 CTCAATTTTCTAGGGGAACTACCCGTGTCTTGGCAAAATTCGACATCCCAACTCC 663
DB 257 CUCAAUUUUUCUAGGGGAGUACCCGUGUGUUCUUGGCAAAAUUCGAGUCCCCAACUUC 316
QY 664 AATCACTCACTCACTCTCTCTCTCACTTCTCTGTTATCTGATGTGTCTGCG 723
DB 317 AAUACACUCACCAACUUCUUCUACAUUUUUCUGUUAUUCGUGAGUUGUUCGCG 376
QY 724 CGTTTATATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 783
DB 377 CUUUUUUUUUUAAAGAGAAAUUCCUGACGAAAGCUUGAUUCUUUUUUUUUUUU 436
QY 784 GACTATCAAGTATGTGCGCGTGTGCTCTTAATTCAGAGATCTTCAACCAACAGACAG 843
DB 437 GAUUAUUAAGAGUAGUUGCCGCUUUUUCUUAUUCGAGAUCAACAAACAGAGUAG 496
QY 844 GGACCAATGAGAGCTGACGACTCTGCTCAAGAACTCTATATCTCTCTCTCTCT 903
DB 497 GGACCAUUGCAAAACUAGCAGACUUCUUGCAAGGCAACUUAAGUUUUCCUAGUUG 556
QY 904 TGTACAAACCTTGGATGGAACAGCTGATATCCATCCATCACTCGGGCTTTC 963
DB 557 UGUACAAACCAACGAGUAGAAUUGACUUGAUUCCAUUCCAGUCCUGGUGUUC 616
QY 964 GGAATAATTCATGGAAGTGGGCTCAGCCGTTTCTCTGAGTCAAGTACAGTACCA 1023
DB 617 GCAAAAUAACUAGAGAGAGGCGCCUACUGCCGUGUUCUUGGUCUACAUUUAACUAGGCCA 676
QY 1024 TTGTTCAGTGTGTTGAGGGCTTTCCGCCCATCTGTTGGCTTTCAGTATATGATGATG 1083
DB 677 UUGUUCACUGUGUUGUAGGCGUUUCCGCCACUGUUGGCUUCAGCUAUAUAGAGUAG 736
QY 1084 TTGTCTGGGGGCCAAGTCTGTACACATCTTGAATCCCTTTTAAACCGCTGTACCAAT 1143
DB 737 UGUUAUUGGGGCGCAAGUCUGACAGCAUUCUGAGUCCCUUAUACCGCUGUACCAAUU 796
QY 1144 TTTCTTGTCTTGGGATATACATTTAAACCTATATAA 1181
DB 797 UUCUUUUUUUUCUGGUAUACAUUUAACCUAACAAA 834
```

```
RESULT 4
US-10-383-317-15
; Sequence 15, Application US/1038317
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha
; APPLICANT: Bass, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/10/383,317
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US/09/247,890
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/074,294
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-11
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/105,509
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 15
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Woodchuck hepatitis B virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(849)
; OTHER INFORMATION: Woodchuck hepatitis B virus (WH8) surface antigen
US-10-383-317-15

Query Match      34.9%; Score 412; DB 6; Length 849;
Best Local Similarity 71.6%; Pred. No. 5.9e-124;
Matches 557; Conservative 0; Mismatches 215; Indels 6; Gaps 1;
```

```
QY 393 TGTGTGGGCTTCAGTTACAGAAACAGTAAACCTGTTCCAGTACTGTCTCCATATC 452
DB 78 TGCCATATGAGATCTTTTACAACTTAAGCCCTCGGTTCTTCTATCTGATCCATATTT 137
QY 453 GTCAATCTTTCGAGATTTGGGAGCCCTGCGGGAACATGAGAACATCACTAGAGATT 512
DB 138 GTCTCTCCCTCGAGACTGGGAGCCCTGCACTGTCAACCGAGATGTACATCAAAATCT 197
QY 513 CCTAGAACCCCTGCTGCTGTGTACAGCGGGGTTTTCTTGTGTGACAAAGATCTTCAAT 572
DB 198 CCTAGAGCTCTCGAGAGATTACAGTGTGATTTCTTGTGACAAAATCTTAACAT 257
QY 573 ACCGAGAGTCTAGACTGTGTGTGACTTCTCAATTTTCTAGGGGGAACCTACCGTGTG 632
DB 258 AGCTAGATATCTAGATTGTGTGTGATCTTCTCAAGTGTTCAGAGGGCATACAGAGTG 317
QY 633 TCTTGGCCAAATTTGGAGTCTCCCAACCTCAATGATCACTACCAACTCTGCTCCCAAC 692
DB 318 CACTGGCCAAATTTGGAGTCTCAATTTGCAAAACCTTGCCAACTCTGCTCAACAC 377
QY 693 TTGTCTGTTATCTGCTGAGATGTCTCGGCGTTTATCAATCTTCTCTCATCTGCT 752
DB 752
```

```
Db      378 TTGCATAGGCTTGTGGATGATCTGCGCGGTTTTATCATATACCTATTAGTCTGCT 437
Qy      753 GCTATGCGCTCATCTTCTTGTGGTCTTCTGAGCATATCAAGTATGTTGCGCGGTTGCTGC 812
Db      438 GCTGTGCGCTCATCTTCTTGTGGTCTTCTGAGCATATCAAGTATGTTGCGCGGTTGCTGC 497
Qy      813 TCTAATTCAGAGATCTTCAACACACAGGAGACATGACAGACTGACAGACTCTGTC 872
Db      498 TCTTCAACCCACAAACAGAAACAGTCA-----ATTGACGACATGACAAATCTCTGC 551
Qy      873 TCAAGGAACTCTATGATATCCCTCTGCTGTGACAAAACCTTGAGTGAATGCTGAC 932
Db      552 ACAAGACATGATATCTCTCTTACTGTGTGTTTAAACCTTACGAGAGAAATTCAC 611
Qy      933 CTGATATCCATCCCATCATCTCTGAGGCTTTCGAAAAATTCCTATGAGAGTGGGCTCAGC 992
Db      612 TTGTGCGCCATCCCTTCATCATAGGCTTTAGAAATTTACTTATGGAAGTGGGCTTAGC 671
Qy      993 CCGTTTCTCTGCTCAGTTACTAGTGCATTTGTTCAGTGTTGCTGAGGCTTTTCCC 1052
Db      672 CCGTTTCTCTGCTCAGTTACTAGTGCATTTGTTCAGTGTTGAGGAAATTTTCCCT 731
Qy      1053 CACTGTTGCTTCACTATATGAGATGATGTTACTGAGGCGGCAAGTCTGTACACAT 1112
Db      732 CATTCGCTGCTTTTGTCTTATATGAGATGATTTGGTTTGGGCGCGCCACTTCTGACAT 791
Qy      1113 CTGAGTCCCTTTTACCGCTGTTACCAATTTCTTTGCTTGGGATACATTTAA 1170
Db      792 CTTACCGCCATTTATCCCATATTTGTTCTGTTTCTTGAATTTGGATATACATTTGA 849
```

```
RESULT 5
PCT-US03-18993-1
; Sequence 1, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolik, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loi, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 128
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-1
```

```
Query Match      10.6%; Score 124.8; DB 1; Length 128;
Best Local Similarity 98.4%; Pred. No. 1.4e-30;
Matches 126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      630 GTGTCTTGGCCAAATTCGAGTCCCAACTCCATCACTACCAACCTCTGCTCTCC 689
Db      1 GTGTCTTGGCCAAATTCGAGTCCCAACTCCATCACTACCAACCTCTGCTCTCC 60
Qy      690 AACTGTCTGCTGATGCTGAGATGCTGCGCGCTTTATCACTTCTCTCTCACTCT 749
Db      61 AATTGTCTGCTGATGCTGAGATGCTGCGCGCTTTATCACTTCTCTCTCACTCT 120
Qy      750 GCTGCTAT 757
Db      121 GCTGCTAT 128
```

```
RESULT 6
PCT-US03-18993-68
; Sequence 68, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolik, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loi, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-68
```

```
Query Match      7.5%; Score 88.2; DB 1; Length 93;
Best Local Similarity 96.8%; Pred. No. 1.1e-18;
Matches 90; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      585 AGACTCGTGTGAGTCTCTCTCAATTTTCTAGGGGAGACTACCGTGTCTTGGCCAAA 644
Db      1 AGACTCGTGTGAGTCTCTCTCAATTTTCTAGGGGAGTACCGGTGTCTTGGCCAAA 60
Qy      645 TTGCGAGTCCCCCACTCTCCATCACTCACCAC 677
Db      61 TTGCGAGTCCCCCACTCTCCATCACTCACCAC 93
```

```
RESULT 7
PCT-US03-18993-69
; Sequence 69, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolik, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loi, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 73
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-69
```

```
Query Match      6.0%; Score 71.4; DB 1; Length 73;
Best Local Similarity 98.6%; Pred. No. 2.9e-13;
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      994 CGTTTCTCTGCTCAGTTACTAGTGCATTTGTTCACTGCTGTTGAGGCTTTCCCCC 1053
Db      1 CGTTTCTCTGCTCAGTTACTAGTGCATTTGTTCACTGCTGTTGAGGCTTTCCCCC 60
```



QY 1054 ACTGTTGGCTTT 1066  
| | | | |  
DB 61 ACTGTTGGCTTT 73

RESULT 8  
PCT-US03-18993-67/c  
; Sequence 67, Application PC/TUS0318993  
; GENERAL INFORMATION:  
; APPLICANT: Linmen, Jeffery M.  
; APPLICANT: Kolk, Daniel P.  
; APPLICANT: Dockter, Janel M.  
; APPLICANT: Getman, Damon K.  
; APPLICANT: Yoshimura, Tadashi  
; APPLICANT: Ho-Sing-Loy, Marcy  
; APPLICANT: Stringfellow, Leslie A.  
; TITLE OF INVENTION: Compositions and Methods for Detecting  
; FILE REFERENCE: GP134-02.UT  
; CURRENT APPLICATION NUMBER: PCT/US03/18993  
; PRIOR FILING DATE: 2003-06-13  
; PRIOR APPLICATION NUMBER: 60/389,393  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 67  
; LENGTH: 71  
; TYPE: DNA  
; ORGANISM: Hepatitis B Virus  
PCT-US03-18993-67

Query Match 5.9%; Score 69.4; DB 1; Length 71;  
Best Local Similarity 98.6%; Pred. No. 1.3e-12;  
Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 736 TTCCTTCATCCTGCTGATGCGCTCATCTTCTTGTTGTTCTTGACATATCAAGT 795  
| | | | |  
DB 71 TTCCTTCATCCTGCTGATGCGCTCATCTTCTTGTTGTTCTTGACATATCAAGT 12

QY 796 ATGTTGCCCGT 806  
| | | | |  
DB 11 ATGTTGCCCGT 1

RESULT 9  
PCT-US03-18993-95/c  
; Sequence 95, Application PC/TUS0318993  
; GENERAL INFORMATION:  
; APPLICANT: Linmen, Jeffery M.  
; APPLICANT: Kolk, Daniel P.  
; APPLICANT: Dockter, Janel M.  
; APPLICANT: Getman, Damon K.  
; APPLICANT: Yoshimura, Tadashi  
; APPLICANT: Ho-Sing-Loy, Marcy  
; APPLICANT: Stringfellow, Leslie A.  
; TITLE OF INVENTION: Compositions and Methods for Detecting  
; FILE REFERENCE: GP134-02.UT  
; CURRENT APPLICATION NUMBER: PCT/US03/18993  
; PRIOR FILING DATE: 2003-06-13  
; PRIOR APPLICATION NUMBER: 60/389,393  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 95  
; LENGTH: 59  
; TYPE: DNA  
; ORGANISM: Hepatitis B Virus  
PCT-US03-18993-95

Query Match 4.9%; Score 57.4; DB 1; Length 59;  
Best Local Similarity 98.3%; Pred. No. 9.7e-09;

Matches 58; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 736 TTCCTTCATCCTGCTGATGCGCTCATCTTCTTGTTGTTCTTGACATATCAAG 794  
| | | | |  
DB 59 TTCCTTCATCCTGCTGATGCGCTCATCTTCTTGTTGTTCTTGACATATCAAG 1

RESULT 10  
US-10-388-918A-7  
; Sequence 7, Application US/10388918A  
; GENERAL INFORMATION:  
; APPLICANT: Browne, Kenneth A.  
; TITLE OF INVENTION: Inversion Probes  
; FILE REFERENCE: GP130-03.CP1  
; CURRENT APPLICATION NUMBER: US/10/388,918A  
; PRIOR FILING DATE: 2003-03-14  
; PRIOR APPLICATION NUMBER: 60/325,600  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: 10/259,272  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 57  
; TYPE: DNA  
; ORGANISM: Hepatitis B Virus  
US-10-388-918A-7

Query Match 4.6%; Score 53.8; DB 6; Length 57;  
Best Local Similarity 96.5%; Pred. No. 1.4e-07;  
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 716 GTCGCGGCGTTTATCATCTTCTTCATCCTGCTGATGCGCTCATCTTCTTGT 772  
| | | | |  
DB 1 GTCGCGGCGTTTATCATCTTCTTCATCCTGCTGATGCGCTCATCTTCTTAT 57

RESULT 11  
PCT-US03-18993-4/c  
; Sequence 4, Application PC/TUS0318993  
; GENERAL INFORMATION:  
; APPLICANT: Linmen, Jeffery M.  
; APPLICANT: Kolk, Daniel P.  
; APPLICANT: Dockter, Janel M.  
; APPLICANT: Getman, Damon K.  
; APPLICANT: Yoshimura, Tadashi  
; APPLICANT: Ho-Sing-Loy, Marcy  
; APPLICANT: Stringfellow, Leslie A.  
; TITLE OF INVENTION: Compositions and Methods for Detecting  
; FILE REFERENCE: GP134-02.UT  
; CURRENT APPLICATION NUMBER: PCT/US03/18993  
; PRIOR FILING DATE: 2003-06-13  
; PRIOR APPLICATION NUMBER: 60/389,393  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 54  
; TYPE: DNA  
; ORGANISM: Hepatitis B Virus  
PCT-US03-18993-4

Query Match 4.3%; Score 50.8; DB 1; Length 54;  
Best Local Similarity 96.3%; Pred. No. 1.3e-06;  
Matches 52; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 768 CTTGTTGTTCTTCGACATATCAAGTATGTCGCCGTTTGTCTCTAATTCC 821  
| | | | |  
DB 54 CTTGTTGTTCTTCGACATATCAAGTATGTCGCCGTTTGTCTCTAATTCC 1

RESULT 12

```
PCT-US03-18993-84
; Sequence 84, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolk, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loy, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 84
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-84
```

```
Query Match          4.2%; Score 50; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1004 GCGTCAGTTACTAGTGCATTTGTCAGTGGTTCGAGGGCTTTCCCCC 1053
DB      1 GCGTCAGTTACTAGTGCATTTGTCAGTGGTTCGAGGGCTTTCCCCC 50
```

```
RESULT 13
PCT-US03-18993-2
; Sequence 2, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolk, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loy, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-2
```

```
Query Match          4.1%; Score 48.4; DB 1; Length 50;
Best Local Similarity 98.0%; Pred. No. 7.8e-06;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      708 CTGATGTCGTGGGCGTTATCATCTCTTCATCTCTGCTGCTAT 757
DB      1 CTGATGTCGTGGGCGTTATCATCATCTCTTCATCTCTGCTGCTAT 50
```

```
RESULT 14
PCT-US03-18993-78
; Sequence 78, Application PC/TUS0318993
; GENERAL INFORMATION:
```

```
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolk, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loy, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 78
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-78
```

```
Query Match          3.6%; Score 43; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      994 CGTTTCTCCTGCTCAGTTACTAGTGCATTTGTCAGTGGT 1036
DB      1 CGTTTCTCCTGCTCAGTTACTAGTGCATTTGTCAGTGGT 43
```

```
RESULT 15
PCT-US03-18993-142/c
; Sequence 142, Application PC/TUS0318993
; GENERAL INFORMATION:
; APPLICANT: Linmen, Jeffery M.
; APPLICANT: Kolk, Daniel P.
; APPLICANT: Dockter, Janel M.
; APPLICANT: Getman, Damon K.
; APPLICANT: Yoshimura, Tadashi
; APPLICANT: Ho-Sing-Loy, Marcy
; APPLICANT: Stringfellow, Leslie A.
; TITLE OF INVENTION: Compositions and Methods for Detecting
; FILE REFERENCE: GP134-02.UT
; CURRENT APPLICATION NUMBER: PCT/US03/18993
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: 60/389,393
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 142
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
PCT-US03-18993-142
```

```
Query Match          3.6%; Score 42.8; DB 1; Length 46;
Best Local Similarity 95.7%; Pred. No. 0.0005;
Matches 44; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      768 CTTGTTGTTCTTCTGACTATCAAGTATGTTGCCGTTGCTT 813
DB      46 CTTATGTTCTTCTGATTATCAAGTATGTTGCCGTTGCTT 1
```

```
Search completed: September 15, 2003, 03:52:06
Job time : 289 secs
```





```

CDS
      /db_xref="GI:1334788"
      /translation="GGLFHLCLILISCSCTVQASXLCLGWL"
      88..639
      /codon_start=1
      /product="core antigen"
      /protein_id="CAA00816.1"
      /db_xref="GI:411874"
      /db_xref="SWISS-PROT:P03147"
      /translation="MDIDPYKEFGATVELSLFSDPFPSVRDLDTAALYRDALES
      PEHGSPIHRLAQLILCWGDMTLATWGNLEDPASDLVSVYNTNVGKFLQLLW
      FHSICLTGRETVLEIVLSFGVWIRTPAYRPNAPILSTLETTVARRGRSPRRRT
      PPSRRRSQSPRRRSQSRRESQC"
      1524..2204
      /codon_start=1
      /product="surface antigen"
      /protein_id="CAA00817.1"
      /db_xref="GI:411875"
      /translation="MENITTSGLFPLVLVQAGFPLLRILITIPQSLDSMTSLNFI
      G
      TVCLGONSQSDPSHNSPTSCPTCPGVRMCLRRFIIPLFLILCLFLVLADYOG
      MLPVPLIPGSSSTSTGSCRTCTPAQGISMSPCCCTPSGNCCTCIPISMAFG
      FLMEWASARFSLVLPFVQWFGVLSPIVMSLVIMMMWGPISLISPLPLPI
      FFLMAYT"

BASE COUNT      638 a      728 c      579 g      798 t
ORIGIN
Query Match      95.9%; Score 1133; DB 6; Length 2743;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY      1 ATGGGCGAGAAATCTTTCACAGCAATCTCTGGAGATTCTTCCGACCAAGTTGAT      60
DB      1035 ATGGGCGAGAAATCTTTCACAGCAATCTCTGGAGATTCTTCCGACCAAGTTGAT      1094
QY      61 CGAGCTTCAGAGCAAAACCAACATCCAGATTGGAGATTCAATCCCAACAGACACC      120
DB      1095 CGAGCTTCAGAGCAAAACCAACATCCAGATTGGAGATTCAATCCCAACAGACACC      1154
QY      121 TGGCCAGACGCCAACAAGGTAGAGCTGGAGCATTCGAGCTGGGGTTCCACCCACCGCAC      180
DB      1155 TGGCCAGACGCCAACAAGGTAGAGCTGGAGCATTCGAGCTGGGGTTCCACCCACCGCAC      1214
QY      181 GGAGGCGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGGATTAACAACACCTTGCACGCAAT      240
DB      1215 GGAGGCGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGGATTAACAACACCTTGCACGCAAT      1274
QY      241 CGGCTCTCTGCTTCCACCAATGCGCAGTGAAGGACGCTTACCCTGCTCTCCACT      300
DB      1275 CGGCTCTCTGCTTCCACCAATGCGCAGTGAAGGACGCTTACCCTGCTCTCCACT      1334
QY      301 TTGAGAAACATCATCTCAAGCCATGAGTGAATCCACAACTTTCCACAAACTCTG      360
DB      1335 CTGAGAACATCATCTCAAGCCATGAGTGAATCCACAACTTTCCACAAACTCTG      1394
QY      361 CAAGATCCAGAGTGAAGGTCTGATATTTCCCTGCTGTGCTCCAGTTCAAGAACAGTA      420
DB      1395 CAAGATCCAGAGTGAAGGTCTGATATTTCCCTGCTGTGCTCCAGTTCAAGAACAGTA      1454
QY      421 AACCTGTTTCCAGTACTGTCTCTCCCATATGTCAATTTTTCAGAGATTGGGACCT      480
DB      1455 AACCTGTTTCCAGTACTGTCTCTCCCATATGTCAATTTTTCAGAGATTGGGACCT      1514
QY      481 GCGGGAACATGAGAACATCATCAGGATTCCTAGAGCCCTGCTGTGTTACAGGGC      540
DB      1515 GCGGGAACATGAGAACATCATCAGGATTCCTAGAGCCCTGCTGTGTTACAGGGC      1574
QY      541 GGGTTTTTCTTGTTGACAAAGATCTCTCAACATACCGACAGTCTAGACTCGTGTGACT      600
DB      1575 GGGTTTTTCTTGTTGACAAAGATCTCTCAACATACCGACAGTCTAGACTCGTGTGACT      1634
QY      601 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTGTCTTGGCAAAATTCGACGTCCCAACC      660
DB      1635 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTGTCTTGGCAAAATTCGACGTCCCAATC      1694

```

```

QY      661 TCCAAATCACTACCAACCTCTGTCTCCCAACTTGTCTCGTGTATCGCTGATGTGTCTG      720
DB      1695 TCCAAATCACTACCAACCTCTGTCTCCCAACTTGTCTCGTGTATCGCTGATGTGTCTG      1754
QY      721 CGGCGTTTATCATTTCTCTTTCATCTCTGCTGTATGCTCTCATCTTCTTGTGTTCTT      780
DB      1755 CGGCGTTTATCATTTCTCTTTCATCTCTGCTGTATGCTCTCATCTTCTTGTGTTCTT      1814
QY      781 CTGACATTCAGAGATATGTCGCGCTGTCTGTCTCAATTCAGAGATCTTCAACCAACAGC      840
DB      1815 CTGACATTCAGAGATATGTCGCGCTGTCTGTCTCAATTCAGAGATCTTCAACCAACAGC      1874
QY      841 ACGGACCATGACAGAGCTGACAGACTCTGCTCAAGAACCTCATGTATCCCTCTGT      900
DB      1875 ACGGACCATGACAGAGCTGACAGACTCTGCTCAAGAACCTCATGTATCCCTCTGT      1934
QY      901 TCTCTACAAAACCTTCGATGAGAACTGACCTGTATTTCCATCCATCATCTCGGCT      960
DB      1935 TCTCTACAAAACCTTCGATGAGAACTGACCTGTATTTCCATCCATCATCTCGGCT      1994
QY      961 TTGGGAAATTCCTATGGAGAGTGGGCTCAGCCGCTTCTCTGCTCAGTTACTAGTG      1020
DB      1995 TTGGGAAATTCCTATGGAGAGTGGGCTCAGCCGCTTCTCTGCTCAGTTACTAGTG      2054
QY      1021 CCATTTGTTCAAGTGTGTGATGAGGCTTCCCACTGTTGGCTTCAATTATGATG      1080
DB      2055 CCATTTGTTCAAGTGTGTGATGAGGCTTCCCACTGTTGGCTTCAATTATGATG      2114
QY      1081 ATGTTGATCGGGGGCCAAAGTCTGTACACCATCTTGAATCCTTTTAAACCGCTTTACCA      1140
DB      2115 ATGTTGATCGGGGGCCAAAGTCTGTACACCATCTTGAATCCTTTTAAACCGCTTTACCA      2174
QY      1141 ATTTCTTTTGTCTTGGGTATACATTTAAACCTTAATAA      1181
DB      2175 ATTTCTTTTGTCTTGGGTATACATTTAAACCTTAATAA      2215

RESULT 3
HPBADYW      2743 bp ss-RNA      linear      VRL 02-AUG-1993
LOCUS      HPBADYW      human hepatitis b virus subtype adyw antigen genes (core antigen
DEFINITION      and surface antigen).
ACCESSION      J02202
VERSION      J02202.1
KEYWORDS      J02202.1 GI:329637
SOURCE      Hepatitis B virus
ORGANISM      Hepatitis B virus
REFERENCE      1 (bases 1 to 2743)
AUTHORS      Paasek, M., Goro, T., Gilbert, W., Zink, B., Schaller, H., Mackay, P.,
      Leadbetter, G., and Murray, K.
TITLE      Hepatitis B virus genes and their expression in E. coli
JOURNAL      Nature 282 (5739), 575-579 (1979)
MEDLINE      81012115
PUBMED      399329
COMMENT      Original source text: hbv subtype adyw from human.
      cf hbvayw and whvayw. hbvayw is hepatitis b core antigen protein,
      and hbvayw is surface antigen protein.
FEATURES
      source
      1..2743
      /organism="Hepatitis B virus"
      /mol_type="genomic RNA"
      /db_xref="taxon:10407"
      88..639
      /note="core antigen"
      /codon_start=1
      /protein_id="AAA45486.1"
      /db_xref="GI:329638"
      /translation="MDIDPYKEFGATVELSLFSDPFPSVRDLDTAALYRDALES
      PEHGSPIHRLAQLILCWGDMTLATWGNLEDPASDLVSVYNTNVGKFLQLLW
      FHSICLTGRETVLEIVLSFGVWIRTPAYRPNAPILSTLETTVARRGRSPRRRT
      PPSRRRSQSPRRRSQSRRESQC"
      1524..2204

```

CDS

```
/note="surface antigen"
/codon_start=1
/db_xref="GI:329639"
/translation="MENITSGFLGLIVLADGFLILRLITIPDLSDMWTSLNPLIG
TYVCLGONSQSPISNHSPTSCPRPCPRMCLRRPIFLPILLCIFLVLIDYOG
MLVCPILPSSSTSTGSCRTCTPAGISYPSCCCTGACDAGCTCIPSSWAEK
FLWEMASRPSMSLVLPFVQWFGSLPIVLSVIMMMWYWSLSVLSLPFLPLPI
FCLMAYI"
BASE COUNT      638 a      728 c      579 g      798 t
ORIGIN      ?
Query Match      95.9%; Score 1133; DB 14; Length 2743;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY      1 ATGGGGAGAAATCTTCCACGCAATCCTGTGGGATCTTCCCGACCAAGTTGGAT 60
      1035 ATGGGGAGAAATCTTCCACGCAATCCTGTGGGATCTTCCCGACCAAGTTGGAT 1094
QY      61 CCAGCCTTCAGAGCAACCAACCAATCCAGATTGGACTTCAATCCCAACAGAGACC 120
      1095 CCAGCCTTCAGAGCAACCAACCAATCCAGATTGGACTTCAATCCCAACAGAGACC 1154
QY      121 TGGCCAGACCCCAACAGTGAAGCTGAGACATTCCGACTGGGGTCCACCCGAC 180
      1155 TGGCCAGACCCCAACAGTGAAGCTGAGACATTCCGACTGGGGTCCACCCGAC 1214
QY      181 GAGAGCCTTTGGGGTGAAGCCCTCAGGCTCAGGGGCAATACAAACCTTCCAGCAAT 240
      1215 GAGAGCCTTTGGGGTGAAGCCCTCAGGCTCAGGGGCAATACAAACCTTCCAGCAAT 1274
QY      241 CCGCCTCTCTCCATCCAAATCCGACATGAGAGAGGAGCCCTACCCGCTGTCTCACT 300
      1275 CCGCCTCTCTCCATCCAAATCCGACATGAGAGAGGAGCCCTACCCGCTGTCTCACT 1334
QY      301 TTGAGAAACCTCACTCTCAAGCCTGAGTGAACCTTCACAACTTTCACCAACTCTG 360
      1335 CTGAGAAACCTCACTCTCAAGCCTGAGTGAACCTTCACAACTTTCACCAACTCTG 1394
QY      361 CAGATCCAGAGAGAGAGTGTATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
      1395 CAGATCCAGAGAGAGAGTGTATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1454
QY      421 AACCTGTTCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGAGATTGGGACCT 480
      1455 AACCTGTTCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGAGATTGGGACCT 1514
QY      481 GCGCGAACAATGAGAAACATCAATCAAGATCTTCTAGAGACCCCTGCTGCTGCTGCTG 540
      1515 GCGCGAACAATGAGAAACATCAATCAAGATCTTCTAGAGACCCCTGCTGCTGCTGCTG 1574
QY      541 GGGTTTTCTTGTGACAGAAATCTCAATACCGAGAGTCTAGACTGTGGTGGACT 600
      1575 GGGTTTTCTTGTGACAGAAATCTCAATACCGAGAGTCTAGACTGTGGTGGACT 1634
QY      601 TCTCTCAATTTCTAGGGGGAACCTACCTGTGTCTTGGCCAAATTTGCGAGTCCCAAC 660
      1635 TCTCTCAATTTCTAGGGGGAACCTACCTGTGTCTTGGCCAAATTTGCGAGTCCCAAC 1694
QY      661 TCCATCACTCAACCACTCTGTCTCTCAATTTGCTGTTATCGCTGAGTGTCTG 720
      1695 TCCATCACTCAACCACTCTGTCTCTCAATTTGCTGTTATCGCTGAGTGTCTG 1754
QY      721 CGGGTTTTTATCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
      1755 CGGGTTTTTATCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814
QY      781 CTGAGACTATCAAGTATGTGCGGTTGTCTCTTAATTCAGAGATTTCAACACGACG 840
      1815 CTGAGACTATCAAGTATGTGCGGTTGTCTCTTAATTCAGAGATTTCAACACGACG 1874
QY      841 ACGGAGCAATGACAGAGCTGACGATCTGCTGCTCAAGAACTCTATGATCTCTCTGT 900
```

```
Db      1875 ACGGAGCTTCGCAAGACCTGCAAGACTCTGCTCAAGAAATCTATGATCTCTCTGT 1934
QY      901 TGCCTGACAAACCTTCGAGTGAACCTGACCTGTATTCATCCATCATCTCTGGAGCT 960
      1935 TGCCTGACAAACCTTCGAGTGAACCTGACCTGTATTCATCCATCATCTCTGGAGCT 1994
QY      961 TTGGAAATTCCTATGAGAGTGGGCTCAGCCCGTTTCTCTGAGTCACTTACTAGTG 1020
      1995 TTGGAAATTCCTATGAGAGTGGGCTCAGCCCGTTTCTCTGAGTCACTTACTAGTG 2054
QY      1021 CCATTTGTTACGTGTTTCTGAGGCTTTCCCCCACTTTGGCTTCAGTTATATGATG 1080
      2055 CCATTTGTTACGTGTTTCTGAGGCTTTCCCCCACTTTGGCTTCAGTTATATGATG 2114
QY      1081 ATGTTGACTGGGGGCAAGTCTGTACCATCTGAGTCCCTTTTACCGCTGTACCA 1140
      2115 ATGTTGACTGGGGGCAAGTCTGTACCATCTGAGTCCCTTTTACCGCTGTACCA 2174
QY      1141 ATTTCTTTTGTCTTGGTATACATTTAAACCTATATAA 1181
      2175 ATTTCTTTTGTCTTGGGATACATTTAAACCTATATAA 2215
Db
RESULT 4
BD181818      8007 bp      DNA      linear      PAT 15-MAY-2003
LOCUS      BD181818
DEFINITION      Hepatitis B virus vectors for gene therapy.
ACCESSION      BD181818
VERSION      BD181818.1 GI:30792736
KEYWORDS      JP 2002320480-A/3.
SOURCE      unidentified
ORGANISM      unidentified
REFERENCE      1 (bases 1 to 8007)
AUTHORS      Ryu,W., Jeong,J.K., Lee,J., Cho,W.Y. and Yoon,G.S.
TITLE      Hepatitis B virus vectors for gene therapy
JOURNAL      Patent: JP 2002320480-A 3 05-NOV-2002;
WANG-SHIK RYU
COMMENT      OS      Unidentified
      PN      JP 2002320480-A/3
      PD      05-NOV-2002
      PF      20-APR-2001 JP 2001122392
      PR      20-APR-2000 KR 2000-21070,12-APR-2001 KR 2001-19645 PI
      WANG-SHIK RYU, JONG KEUN JEONG, JEHAN LEE, MOO YOUNG CHO, GYE PI
      SOON YOUNG
      PC      C12N15/09, A61K35/74, A61K48/00, A61P31/20, C12N7/00, C12N15/00 CC
      PCMV-HBV/30 Pull Sequences
      CC      8007 bp ms-DNA circular
      CC      From HBV-ayw
      CC      #1 : #1820 of HBV-ayw (accession number J02203) CC
      HBV-ayw (178 bp overlapping the HBV genome-length) CC #1 - 3360 ;
      secondary structure: bases 30 - 90 CC 3' - epsilon secondary
      structure ; bases 3212 - 3272 CC 5' - DR1 ; bases 7 - 17
      CC 3' - DR1 ; bases 3189 - 3199
      CC Poly A signal ; bases 3281-3286
      CC 3' - DR2 ; bases 2955 - 2965
      CC Core ORF ; bases 84 - 632 (exclude stop codon) CC S1 ORF ; bases
      ORF ; bases 490 - 2985 (exclude stop codon) CC S2 ORF ; bases 1355 - 2197
      1031 - 2197 (exclude stop codon) CC S ORF ; bases 1520 - 2197 (exclude stop
      codon) CC X ORF ; bases 2739 - 3200 (exclude stop codon) CC
      From pcDNA1/Amp
      CC Col E1 origin ; bases 5103-5689 (1-587 of pcDNA1/Amp) CC M13
      origin ; bases 5690-6282
      CC Ampicillin gene ; bases 6462-7405
      CC CMV promoter ; bases 7406-7999
      CC SP6 primer sequence ; bases 3372-3390
      CC Splice and polyA ; bases 3391-4089
      FH      Key      Location/Qualifiers
      FT      source      1..8007
      /organism='Unidentified'.
```

FEATURES  
source Location/Qualifiers  
1..8007  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
BASE COUNT 2022 a 1980 c 1824 g 2181 t  
ORIGIN

Query Match 95.5%; Score 1128.2; DB 6; Length 8007;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1148; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGGGAGAAATCTTTCACAGCAATCTCTGGATTCTTCCGACCAAGTTGAT 60  
DB 1031 ATGGGAGAAATCTTTCACAGCAATCTCTGGATTCTTCCGACCAAGTTGAT 1090

QY 61 CAGGCTTCAGAGCAAAACCAATCAATGGGACTTCAATCCCAAGAGACCC 120  
DB 1091 CAGGCTTCAGAGCAAAACCAATCAATGGGACTTCAATCCCAAGAGACCC 1150

QY 121 TGGCCAGACGCCAACAAAGGTAGAGCTGAGCAATTCGGATCGGGCTTCA 180  
DB 1151 TGGCCAGACGCCAACAAAGGTAGAGCTGAGCAATTCGGATCGGGCTTCA 1210

QY 181 GAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGCAATACAACTTGGCAGCA 240  
DB 1211 GAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGCAATACAACTTGGCAGCA 1270

QY 241 CGGCTCTCTGCTTCCACCAATGCGCAGTCAAGAAAGGACCTTACCCGCT 300  
DB 1271 CGGCTCTCTGCTTCCACCAATGCGCAGTCAAGAAAGGACCTTACCCGCT 1330

QY 301 TTGAGAAACACTCATCTCCTCAAGCAATGAGTGAAGTCCACAATTTCCACA 360  
DB 1331 TTGAGAAACACTCATCTCCTCAAGCAATGAGTGAAGTCCACAATTTCCACA 1390

QY 361 CAAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGTGGCTCCAGTTCAAGAA 420  
DB 1391 CAAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGTGGCTCCAGTTCAAGAA 1450

QY 421 AACCTGTTCGACTCATCTCTCTCCATATGTCATTTTTCGAGATTGGGACCT 480  
DB 1451 AACCTGTTCGACTCATCTCTCTCCATATGTCATTTTTCGAGATTGGGACCT 1510

QY 481 GGGGGAACATGGAAGCAATCATCAGGATTCCTAGGACCCCTGCTCTTACAG 540  
DB 1511 GGGGGAACATGGAAGCAATCATCAGGATTCCTAGGACCCCTGCTCTTACAG 1570

QY 541 GGGTTTTCTTGTGACAGAAATCCATACAGATCGCAGATCGTGTGAGT 600  
DB 1571 GGGTTTTCTTGTGACAGAAATCCATACAGATCGCAGATCGTGTGAGT 1630

QY 601 TCTCTCAATTTTCTAGGGGGAATCAACGCTGTCTTGGCCAAATTCGAGTCC 660  
DB 1631 TCTCTCAATTTTCTAGGGGGAATCAACGCTGTCTTGGCCAAATTCGAGTCC 1690

QY 661 TCCATTCATCTCAACCAACCTCTCTCTCTCAATTCCTGATGATGTGTCTG 720  
DB 1691 TCCATTCATCTCAACCAACCTCTCTCTCTCAATTCCTGATGATGTGTCTG 1750

QY 721 CGGCTTTTATCATCTTCTCTCTCAATTCCTGATGATGTGTCTG 780  
DB 1751 CGGCTTTTATCATCTTCTCTCTCAATTCCTGATGATGTGTCTG 1810

QY 781 CTGAATCATCAAGTATGTGCGGCTTGTCTTAAATTCAGATCTTCAACCAAC 840  
DB 1811 CTGAATCATCAAGTATGTGCGGCTTGTCTTAAATTCAGATCTTCAACCAAC 1870

QY 841 ACGGAGCAATGAGAGCTGCAAGATCTGCTCAAGAACTCTATGATCCCTCTCT 900  
DB 1871 ACGGAGCAATGAGAGCTGCAAGATCTGCTCAAGAACTCTATGATCCCTCTCT 1930

QY 901 TGCCTACAAAACCTTCGATGGAACCTGATATTCATCCATCATCTGGGCT 960

DB 1931 TGCCTACAAAACCTTCGAGCAAGATGACCTGATATCCATCATCTGGGCT 1990

QY 961 TTGGGAAATTCCTATGGAAGTGGGCTCAGCCCTTCTCTGCTCAGTTACTAG 1020  
DB 1991 TTGGGAAATTCCTATGGAAGTGGGCTCAGCCCTTCTCTGCTCAGTTACTAG 2050

QY 1021 CCATTTGTTAGTGGTGGTGAAGGCTTCCCACTGTTTGGCTTCAATATGATG 1080  
DB 2051 CCATTTGTTAGTGGTGGTGAAGGCTTCCCACTGTTTGGCTTCAATATGATG 2110

QY 1081 ATGTTGATCGGAGGCCAAGTCTGACACCAATCTTGAATCCCTTTTACCG 1140  
DB 2111 ATGTTGATCGGAGGCCAAGTCTGACACCAATCTTGAATCCCTTTTACCG 2170

QY 1141 ATTTTCTTTTGTCTTTGGTATACATTTTAACTTAACTTAA 1181  
DB 2171 ATTTTCTTTTGTCTTTGGTATACATTTTAACTTAACTTAA 2211

RESULT 5  
AR011346 1201 bp DNA linear PART 04-DEC-1998  
LOCUS AR011346  
DEFINITION Sequence 218 from patent US 5762938.  
ACCESSION AR011346  
VERSION AR011346.1 GI:3969336  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1201)  
Paoletti,E., Pertus,M.E., Taylor,J., Tartaglia,J., Norton,E.K.,  
Riviere,M., de Taisne,C., Limbach,K.J., Johnson,G.P., Pincus,S.E.,  
Cox,W.I., Audomest,D.-C.,Francis, and Gettig,R.Robert.  
Modified recombinant vaccinia virus and expression vectors thereof  
Patent: US(5762938-A)218 09-JUN-1998;  
TITLE JOURNAL  
FEATURES  
source Location/Qualifiers  
1..1201  
/organism="unknown"  
BASE COUNT 265 a 350 c 255 g 331 t  
ORIGIN

Query Match 94.6%; Score 1117.2; DB 6; Length 1201;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGGGAGAAATCTTTCACAGCAATCTCTGGATTCTTCCGACCAAGTTGAT 60  
DB 32 ATGGGAGAAATCTTTCACAGCAATCTCTGGATTCTTCCGACCAAGTTGAT 91

QY 61 CAGGCTTCAGAGCAAAACCAATCAATGGGACTTCAATCCCAAGAGACCC 120  
DB 92 CAGGCTTCAGAGCAAAACCAATCAATGGGACTTCAATCCCAAGAGACCC 151

QY 121 TGGCCAGACGCCAACAAAGGTAGAGCTGAGCAATTCGGATCGGGCTTCA 180  
DB 152 TGGCCAGACGCCAACAAAGGTAGAGCTGAGCAATTCGGATCGGGCTTCA 211

QY 181 GAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGCAATACCAAACTTGCAGCA 240  
DB 212 GAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGCAATACCAAACTTGCAGCA 271

QY 241 CGGCTCTCTGCTTCCACCAATGCGCAGTCAAGAAAGGACCTTACCCGCT 300  
DB 272 CGGCTCTCTGCTTCCACCAATGCGCAGTCAAGAAAGGACCTTACCCGCT 331

QY 301 TTGAGAAACACTCATCTCCTCAAGCAATGAGTGAAGTCCACAATTTTCAAC 360  
DB 332 TTGAGAAACACTCATCTCCTCAAGCAATGAGTGAAGTCCACAATTTTCAAC 391

QY 361 CAAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGTGAGCTCCAGTTCAAGAA 420  
DB 392 CAAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGTGAGCTCCAGTTCAAGAA 451

QY 421 AACCTGTTCGACTACTGTCTCTCCCATATGTCATCTTTCTGAGAGATTGGGACCT 480  
DB 452 AACCTGTTCGACTACTGTCTCTCCCATATGTCATCTTTCTGAGAGATTGGGACCT 511  
QY 481 GCGGGAAACATGGAACATCATCATAGAGATTCTTAGAGACCCGCTCGTGTTCAGAGC 540  
DB 512 GCGGTAAACATGGAACATCATCATAGAGATTCTTAGAGACCCGCTCGTGTTCAGAGC 571  
QY 541 GGGTTTTCTGTGTGACAGAGATCTCAATACCGAGAGTCTAGACTCTGTGTGACT 600  
DB 572 GGGTTCTCTGTGTGACAGAGATCTCAATACCGAGAGTCTAGACTCTGTGTGACT 631  
QY 601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAATTCGACGTCCCAAC 660  
DB 632 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAATTCGACGTCCCAAC 691  
QY 661 TCCAAATACATCCAACTCTCTGTCTCCAACTTGTCTGTGTATGCTGTGTGTGTG 720  
DB 692 TCCAAATACATCCAACTCTCTGTCTCCAACTTGTCTGTGTATGCTGTGTGTGTG 751  
QY 721 CGGCGTTTATCATCTCTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
DB 752 CGGCGTTTATCATCTCTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 811  
QY 781 CTGACTATCAAGGATGTGTGCGGTTGTCTCTAAATTCAGAGATCTTCAACCCAGC 840  
DB 812 CTGACTATCAAGGATGTGTGCGGTTGTCTCTAAATTCAGAGATCTTCAACCCAGC 871  
QY 841 ACCGGACATGACAGAGCTGACAGCTCTGTCTCAAGAACTCTATGTATCCCTCTGT 900  
DB 872 ACCGGACATGACAGAGCTGACAGCTCTGTCTCAAGAACTCTATGTATCCCTCTGT 931  
QY 901 TGTGTACAAACCTTGTGATGAACTGACACTGTATTCCTCATCCATCATCTCTGGCT 960  
DB 932 TGTGTACAAACCTTGTGATGAACTGACACTGTATTCCTCATCCATCATCTCTGGCT 991  
QY 961 TTGGGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCACTTACTAGTG 1020  
DB 992 TTGGGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCACTTACTAGTG 1051  
QY 1021 CCAATTGTTCAGTGTCTGAGAGGCTTCCCCCACTGTGTGGCTTCAAGTATATGAGTG 1080  
DB 1052 CCAATTGTTCAGTGTCTGAGAGGCTTCCCCCACTGTGTGGCTTCAAGTATATGAGTG 1111  
QY 1081 ATGTTGTATGAGGAGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTACCA 1140  
DB 1112 ATGTTGTATGAGGAGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTACCA 1171  
QY 1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170  
DB 1172 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1201

RESULT 6  
11784 11784 1201 bp DNA linear PAT 07-OCT-1996  
LOCUS Sequence 218 from patent US 5494807.  
DEFINITION 11784  
ACCESSION 11784  
VERSION 11784.1 GI:1598339  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclashed.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Paolucci, E., Perkins, M. E., Taylor, J., Tartaglia, J., Norton, E. K.,  
Riviere, M., de Taisne, C., Limbach, K. J., Johnson, G. P., Pincus, S. E.,  
Cox, W. I., Audommet, J.-C. F., and Getlig, R. R.  
TITLE NVNAC vaccinia virus recombinants comprising heterologous inserts  
JOURNAL Patent: US 5494807-A 218 27-FEB-1996;  
FEATURES  
location/Qualifiers  
1..1201  
/organism="unknown"

BASE COUNT 265 a 350 c 255 g 331 t  
ORIGIN  
Query Match 94.6%; Score 1117.2; DB 6; Length 1201;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 1 ATGGGGGAGAAATCTTTCACACCAATCTCTGGGAATCTTTCGCCACACAGTTGGAT 60  
DB 32 ATGGGGGAGAAATCTTTCACACCAATCTCTGGGAATCTTTCGCCACACAGTTGGAT 91  
QY 61 CCAAGCTTCAGACCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGGACACC 120  
DB 92 CCAAGCTTCAGACCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGGACACC 151  
QY 121 TGAGCAGACCCCAACCAAGTAGAGCTGAGCAATTCGSACTGGGGTTTCAACCCACGAC 180  
DB 152 TGAGCAGACCCCAACCAAGTAGAGCTGAGCAATTCGSACTGGGGTTTCAACCCACGAC 211  
QY 181 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACACAACTTTGCCAGCAAT 240  
DB 212 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACACAACTTTGCCAGCAAT 271  
QY 241 CCGCTTCCTGCTTCACCAATGGCCAGTAGAGAGGAGAGGAGAGGAGGAGGAGGAGG 300  
DB 272 CCGCTTCCTGCTTCACCAATGGCCAGTAGAGAGGAGAGGAGGAGGAGGAGGAGG 331  
QY 301 TTGAGAAACATCATCTCTCAAGCAATGACAGTGGAACTCCACAACTTTCACCAACTCTG 360  
DB 332 TTGAGAAACATCATCTCTCAAGCAATGACAGTGGAACTCCACAACTTTCACCAACTCTG 391  
QY 361 CAAGATCCAGAGTGAAGGCTGTGTATTTCTCTGTGTGTGTGTGTGTGTGTGTGTGT 420  
DB 392 CAAGATCCAGAGTGAAGGCTGTGTATTTCTCTGTGTGTGTGTGTGTGTGTGTGTGT 451  
QY 421 AACCTGTTCGACTACTGTCTCTCCCATATGTCATCTTTCTGAGAGATTGGGACCT 480  
DB 452 AACCTGTTCGACTACTGTCTCTCCCATATGTCATCTTTCTGAGAGATTGGGACCT 511  
QY 481 GCGGGAAACATGGAACATCATCATAGAGATTCTTAGAGACCCGCTCGTGTTCAGAGC 540  
DB 512 GCGGTAAACATGGAACATCATCATAGAGATTCTTAGAGACCCGCTCGTGTTCAGAGC 571  
QY 541 GGGTTTTCTGTGTGACAGAGATCTCAATACCGAGAGTCTAGACTCTGTGTGACT 600  
DB 572 GGGTTCTCTGTGTGACAGAGATCTCAATACCGAGAGTCTAGACTCTGTGTGACT 631  
QY 601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAATTCGACGTCCCAAC 660  
DB 632 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAATTCGACGTCCCAAC 691  
QY 661 TCCAAATACATCCAACTCTCTGTCTCCAACTTGTCTGTGTATGCTGTGTGTGTG 720  
DB 692 TCCAAATACATCCAACTCTCTGTCTCCAACTTGTCTGTGTATGCTGTGTGTGTG 751  
QY 781 CTGACTATCAAGGATGTGTGCGGTTGTCTCTAAATTCAGAGATCTTCAACCCAGC 840  
DB 812 CTGACTATCAAGGATGTGTGCGGTTGTCTCTAAATTCAGAGATCTTCAACCCAGC 871  
QY 841 ACCGGACATGACAGAGCTGACAGCTCTGTCTCAAGAACTCTATGTATCCCTCTGT 900  
DB 872 ACCGGACATGACAGAGCTGACAGCTCTGTCTCAAGAACTCTATGTATCCCTCTGT 931  
QY 901 TGTGTACAAACCTTGTGATGAACTGACACTGTATTCCTCATCCATCATCTCTGGCT 960  
DB 932 TGTGTACAAACCTTGTGATGAACTGACACTGTATTCCTCATCCATCATCTCTGGCT 991  
QY 961 TTGGGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCACTTACTAGTG 1020



Db 992 TTCGGAATAATCTTATGGAGTGGGCTTCAGCCCGTTTCTCTGCTAGTTACTAGTG 1051  
Qy 1021 CCATTGTTCACTGTTCTGTAAGGCTTTTCCCACTGTTGGCTTTAGTTATATGATG 1080  
Db 1052 CCATTGTTCACTGTTCTGTAAGGCTTTTCCCACTGTTGGCTTTAGTTATATGATG 1111  
Qy 1081 ATGTGTATGGGGGCAAGTCTGTACACCAATCTTGAGTCCCTTTTACCGCTGTTACCA 1140  
Db 1112 ATGTGTATGGGGGCAAGTCTGTACACCAATCTTGAGTCCCTTTTACCGCTGTTACCA 1171  
Qy 1141 ATTTCTTTTGTCTTGGGTATACATTTAA 1170  
Db 1172 ATTTCTTTTGTCTTGGGTATACATTTAA 1201

## RESULT 7

LOCUS AR011345 1285 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 215 from patent US 5762938.  
ACCESSION AR011345  
VERSION AR011345.1 GI:3969335  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

## REFERENCE

AUTHORS Paolietti,E., Perkins,M.E., Taylor,J., Tartaglia,J., Norton,E.K.,  
Riviere,M., de Taisne,C., Limbach,K.J., Johnson,G.P., Pincus,S.E.,  
Cox,W.I., Audemmet,J.-C., Francis, and Gettlig,R.Robert.  
TITLE Modified recombinant vaccinia virus and expression vectors thereof  
JOURNAL Patent: US 5762938-A 215 09-JUN-1998;  
FEATURES Location/Qualifiers  
source 1..1285  
/organism="unknown"

BASE COUNT 284 a 364 c 272 g 365 t  
ORIGIN

Query Match 94.6%; Score 1117.2; DB 6; Length 1285;  
Best Local Similarity 97.2%; Pred. No. 0;

Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ATGGGAGCAATCTTTCACGACCAATCTCTGGAGTTCTTCCGACCAAGTTGAT 60  
Db 116 ATGGGAGCAATCTTTCACGACCAATCTCTGGAGTTCTTCCGACCAAGTTGAT 175  
Qy 61 CGAGCTTTCAGAGCAACCAACATCCAGATTGGGACTTCAATCCCAAGAGACAC 120  
Db 176 CGAGCTTTCAGAGCAACCAACATCCAGATTGGGACTTCAATCCCAAGAGACAC 235  
Qy 121 TGGCCAGAGCGCAACAAAGGTAGAGCTGAGCACTTGGAGCTGGGGTTTCACCCAACGAC 180  
Db 236 TGGCCAGAGCGCAACAAAGGTAGAGCTGAGCACTTGGAGCTGGGGTTTCACCCAACGAC 295  
Qy 181 GGAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGGCAACAAACCTTGCAGCAAT 240  
Db 296 GGAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGGCAATCAAACTTTGCAGCAAT 355  
Qy 241 CGGCTCTCTGCTTTCACCAATGCGCAGTCAAGAGGACCTTACCCCGCTGTCTCACT 300  
Db 356 CGGCTCTCTGCTTTCACCAATGCGCAGTCAAGAGGACCTTACCCCGCTGTCTCACT 415  
Qy 301 TTGAGAAACACTCATCTCCTCAAGCATGAGTGAATCTCCACAATTTCCACCAAACTTG 360  
Db 416 TTGAGAAACACTCATCTCCTCAAGCATGAGTGAATTTCCACAACCTTTTACCAAACTTG 475  
Qy 361 CAAGATCCAGAGTGAAGGTCTGTATTTCCCTGCTGTGGCTCCAGTTCAAGAAACAGTA 420  
Db 476 CAAGATCCAGAGTGAAGGTCTGTATTTCCCTGCTGTGGCTCCAGTTCAAGAAACAGTA 535  
Qy 421 AACCTGTTCGAGTACTGTCTCTCCCAATGTCTCAATCTTTCGAGAGATTGGGACCT 480  
Db 536 AACCTGTTCGAGTACTGTCTCTCCCAATGTCTCAATCTTTCGAGAGATTGGGACCT 595

Qy 481 GCGCGAATCATGAGAAACATCATCAGATTCTAGAACCCCTGCTGCTTTACAGCG 540  
Db 596 GCGTGAACATGAGAAACATCATCAGATTCTAGAACCCCTGCTGCTTTACAGCG 655  
Qy 541 GGGTTTTCTTGTGACAAAGATCTCTACATACCGCAGAGTCTAGTCTGGTGGACT 600  
Db 656 GGGTTTTCTTGTGACAAAGATCTCTACATACCGCAGAGTCTAGTCTGGTGGACT 715  
Qy 601 TCTCAATTTCTTGGGGGAATCTACCGTGTCTTGGCAAAATTCGAGTCCCAAC 660  
Db 716 TCTCAATTTCTTGGGGGAATCTACCGTGTCTTGGCAAAATTCGAGTCCCAAC 775  
Qy 721 CGGCTTTTATCATCTTCTCTTCAATCTGCTCTATGCTTATCTTGTGGTCTT 780  
Db 836 CGGCTTTTATCATCTTCTCTTCAATCTGCTCTATGCTTATCTTGTGGTCTT 895  
Qy 781 CTGACTATCAAGTATGTTGGCGGTTTGTCTCTAATTCAGAGATCTTCAACACAC 840  
Db 896 CTGACTATCAAGTATGTTGGCGGTTTGTCTCTAATTCAGAGATCTTCAACACAC 955  
Qy 841 ACGGAGCATGAGAGCGCTGACAGACTCTCTCAAGAACTCTATGATCCCTCTGT 900  
Db 956 ACGGAGCATGAGAGCGCTGACAGACTCTCTCAAGAACTCTATGATCCCTCTGT 1015  
Qy 901 TCTGTACAAAACCTTGGAGTGAATCTGACCTGTATTTCCATCCATCTCTGGCT 960  
Db 1016 TCTGTACAAAACCTTGGAGTGAATCTGACCTGTATTTCCATCCATCTCTGGCT 1075  
Qy 961 TTGGAAATTCCTATGGAGTGGGCGTCAAGCGGTTTCTCTGCTCAAGTTACTAGTG 1020  
Db 1076 TTGGAAATTCCTATGGAGTGGGCGTCAAGCGGTTTCTCTGCTCAAGTTACTAGTG 1135  
Qy 1021 CCATTGTTCACTGTTCTGTAAGGCTTTTCCCACTGTTGGCTTTAGTTATATGATG 1080  
Db 1136 CCATTGTTCACTGTTCTGTAAGGCTTTTCCCACTGTTGGCTTTAGTTATATGATG 1195  
Qy 1081 ATGTGTATGGGGGCAAGTCTGTACACCAATCTTGAGTCCCTTTTACCGCTGTTACA 1140  
Db 1196 ATGTGTATGGGGGCAAGTCTGTACACCAATCTTGAGTCCCTTTTACCGCTGTTACA 1255  
Qy 1141 ATTTCTTTTGTCTTGGGTATACATTTAA 1170  
Db 1256 ATTTCTTTTGTCTTGGGTATACATTTAA 1285

## RESULT 8

LOCUS 117983 1285 bp DNA linear PAT 07-OCT-1996  
DEFINITION Sequence 215 from patent US 5494807.  
ACCESSION 117983  
VERSION 117983.1 GI:1598338  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

## REFERENCE

AUTHORS Paolietti,E., Perkins,M.E., Taylor,J., Tartaglia,J., Norton,E.K.,  
Riviere,M., de Taisne,C., Limbach,K.J., Johnson,G.P., Pincus,S.E.,  
Cox,W.I., Audemmet,J.-C., F. and Gettlig,R.R.  
TITLE NYVAC vaccinia virus recombinants comprising heterologous inserts  
JOURNAL Patent: US 5494807-A 215 27-FEB-1996;  
FEATURES Location/Qualifiers  
source 1..1285  
/organism="unknown"

BASE COUNT 284 a 364 c 272 g 365 t  
ORIGIN

Query Match 94.6%; Score 1117.2; DB 6; Length 1285;  
Best Local Similarity 97.2%; Pred. No. 0;

Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;				
Qy	1	ATGGGAGAGATCTTTCCACGAGATCTCTGGGATTTCTTCCGACCAAGTTGGAT	60	
Db	116	ATGGGAGAGATCTTTCCACGAGATCTCTGGGATTTCTTCCGACCAAGTTGGAT	175	
Qy	61	CCAGCTTCAGAGCAACACCAACATCCAGATTGGGACTTCGAATCCCAACAGACACC	120	
Db	176	CCAGCTTCAGAGCAACACCAACATCCAGATTGGGACTTCGAATCCCAACAGACACC	235	
Qy	121	TTGGCAACCGCCAAAGAGTAGAGCTGGAGCATTCGAGCTGGGGTTCACCCCGCGAC	180	
Db	236	TTGGCAACCGCCAAAGAGTAGAGCTGGAGCATTCGAGCTGGGGTTCACCCCGCGAC	295	
Qy	181	GSAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATACAAACCTTCCAGCAAT	240	
Db	296	GSAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATACAAACCTTCCAGCAAT	355	
Qy	241	CCGCTCTCTGCTTCCACCAATCGCCAGTCAGGAAGGAGCCCTACCCGCTGTCCACT	300	
Db	356	CCGCTCTCTGCTTCCACCAATCGCCAGTCAGGAAGGAGCCCTACCCGCTGTCCACT	415	
Qy	301	TTGGAACAACCTCATCTCCCAAGCCATGAGTGAACCTCCAACTTTCACCAACTCTG	360	
Db	416	TTGGAACAACCTCATCTCCCAAGCCATGAGTGAACCTCCAACTTTCACCAACTCTG	475	
Qy	361	CAAGATCCAGAGTAGAGGTCTGTATTTCCCTGCTGGTGGCTCAGTTCCAGAACAGTA	420	
Db	476	CAAGATCCAGAGTAGAGGTCTGTATTTCCCTGCTGGTGGCTCAGTTCCAGAACAGTA	535	
Qy	421	AACTCTGTTCCGACTACTGTCTTCCCATATGTCATCTTTCTGAGAGATTGGGACCT	480	
Db	536	AACTCTGTTCCGACTACTGTCTTCCCATATGTCATCTTTCTGAGAGATTGGGACCT	595	
Qy	481	GGCGGGAACATGAGGAACATCAATCAAGGATTCCTAGGAGCCCGCTCGTGTCAAGCG	540	
Db	596	GGCGGGAACATGAGGAACATCAATCAAGGATTCCTAGGAGCCCGCTCGTGTCAAGCG	655	
Qy	541	GGGTTTTTCTGTGACAAAGATCTCACAATCCGAGAGTCTAGACTCGTGTGAGCT	600	
Db	656	GGGTTTTTCTGTGACAAAGATCTCACAATCCGAGAGTCTAGACTCGTGTGAGCT	715	
Qy	601	TCTCTCAATTTTCTAGGGGGAATACCGTGTGTCTTGGCCAAAATTGGCAGTCCCCAAC	660	
Db	716	TCTCTCAATTTTCTAGGGGGAATACCGTGTGTCTTGGCCAAAATTGGCAGTCCCCAAC	775	
Qy	661	TCCAAATCACTCAACCACTCTGTCTCCCACTTGTCTGTTATTCGTGGATGTCGTG	720	
Db	776	TCCAAATCACTCAACCACTCTGTCTCCCACTTGTCTGTTATTCGTGGATGTCGTG	835	
Qy	721	CGGCGTTTATCATCTTCTCTTCACTCTGCTGCTATGCTCATCTTCTTGTGGTCTT	780	
Db	836	CGGCGTTTATCATCTTCTCTTCACTCTGCTGCTATGCTCATCTTCTTGTGGTCTT	895	
Qy	781	CTGAGCTATCAAGGATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAACAC	840	
Db	896	CTGAGCTATCAAGGATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAACAC	955	
Qy	841	ACGGGACCATGCGAGGCTGACAGCTCTGTCTCAAGGAACTCTATGATTCCTCCCTG	900	
Db	956	ACGGGACCATGCGAGGCTGACAGCTCTGTCTCAAGGAACTCTATGATTCCTCCCTG	1015	
Qy	901	TGCTGTACAAAACCTTCGAGTGAATACTGACCTGTATTCCTATCCATCCATCCGGGCT	960	
Db	1016	TGCTGTACAAAACCTTCGAGTGAATACTGACCTGTATTCCTATCCATCCATCCGGGCT	1075	
Qy	961	TTTCGAAAATTCCTATGGAAGTGGGCTCAGCCGTTTCTCTGGCTCAGTTTACTAGTG	1020	
Db	1076	TTTCGAAAATTCCTATGGAAGTGGGCTCAGCCGTTTCTCTGGCTCAGTTTACTAGTG	1135	
Qy	1021	CCATTGTTCAGTGGTGTGTAGGGGTTTCCCACTGTTGAGCTTTCAGTTATATGATG	1080	
Db	1136	CCATTGTTCAGTGGTGTGTAGGGGTTTCCCACTGTTGAGCTTTCAGTTATATGATG	1195	

Qy	1081	ATGTTTACTGGGGGCCCAAGTCTGTACACCATCTTAGTCCCTTTTACCGCTGTACCA	1140
Db	1196	ATGTGATATGGGGGCCCAAGTCTGTACACCATCTTAGTCCCTTTTACCGCTGTACCA	1255
Qy	1141	ATTTTCTTTTGTCTTTGGGTATACATTAA	1170
Db	1256	ATTTTCTTTTGTCTTTGGGTATACATTAA	1285
RESULT 9			
A32618	A32618	2342 bp	DNA
LOCUS	A32618		linear
DEFINITION	Synthetic coding seq for pre-S1,pre-S2 and S hepatitis B antigen peptides.		
ACCESSION	A32618		
VERSION	A32618.1	GI:1567467	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1 (bases 1 to 2342)		
AUTHORS	Thoma,H.A.D.		
TITLE	Peptide comprising hepatitis B surface antigen		
JOURNAL	Patent: EP 0304578-A 1 01-MAR-1989; Eptec Limited		
FEATURES	Location/Qualifiers		
source	1..2342		
	/organism="synthetic construct"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32630"		
BASE COUNT	493 a	661 c	524 g 664 t
ORIGIN			
Query Match 94.6%; Score 1117.2; DB 6; Length 2342; Best Local Similarity 97.1%; Pred. No. 0; Matches 1148; Conservative 0; Mismatches 33; Indels 1; Gaps 1;			
Qy	1	ATGGGGAGAGATCTTTCACCGAGATCTCTGGGATTTCTTCCGACCAAGTTGGAT	60
Db	12	ATGGGGAGAGATCTTTCACCGAGATCTCTGGGATTTCTTCCGACCAAGTTGGAT	71
Qy	61	CCAGCTTCAGAGCAACACCAACATCCAGATTGGGACTTCAATCCCAACAGACACC	120
Db	72	CCAGCTTCAGAGCAACACCGCAATCCAGATTGGGACTTCAATCCCAACAGACACC	131
Qy	121	TTGGCCAGACGCCAACAGGTAGAGCTGGAGCATTTCCGATGGGGTTCACCCACCGCAC	180
Db	132	TTGGCCAGACGCCCAACAGGTAGAGCTGGAGCATTTGGGCTTTCACCCACCGCAC	191
Qy	181	GGAGGCTCTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATACACAACCTTCCAGCAAT	240
Db	192	GGAGGCTCTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATACACAACCTTCCAGCAAT	251
Qy	241	CCGCTCTCTGCTTCCACCAATCGCCAGTCAGGAAGGAGGAGCTTACCCGCTGTCCACT	300
Db	252	CCGCTCTCTGCTTCCACCAATCGCCAGTCAGGAAGGAGGAGCTTACCCGCTGTCCACT	311
Qy	301	TTGGAACAACATCATCTCCCAAGCCATGAGTGAACCTCCAACTTTCACCAAACTCTG	360
Db	312	TTGGAACAACATCATCTCCAGGCCATGAGTGAATTCACCAACCTTTCACCAAACTCTG	371
Qy	361	CAAGATCCCAAGTAGAGGTCTGTATTTCCCTGCTGGTGGCTCAGTTGAGAAACAGTA	420
Db	372	CAAGATCCCAAGTAGAGGTCTGTATTTCCCTGCTGGTGGCTCAGTTGAGAAACAGTA	431
Qy	421	AACTCTGTTCCGACTACTGTCTCTCCCATATGTCATCTTTCTGAGAGATTGGGACCT	480
Db	432	AACTCTGTTCCGACTACTGTCTCTCCCATATGTCATCTTTCTGAGAGATTGGGACCT	491
Qy	481	GGCGGGAACATGAGGAACATCACATCGAATTCCTAGAGAACCCGCTCGTGTACAGCG	540
Db	492	GGCGGGAACATGAGGAACATCACATCGAATTCCTAGAGAACCCCTTCGTGTGTACAGCG	551

OY		541	GCGTTTTTCTGTTACAAAGATCTCCCAATAACGGCAGAGCTTGACTCGGGGTGA	600
Db		552	GGGTTTTTCTTGTCACAAGAATCTCCCAATACCGGCAGAGCTTGACTCGGGGTGA	611
OY		601	TCTCTCAATTTCCTAGGGGAAACTAACGGTGTCTCTTGCCCAAATAATTCGACAGTCCC	660
Db		612	TCTCTCAATTTCCTAGGGGAAACTAACGGTGTCTCTTGCCCAAATAATTCGACAGTCCC	671
OY		661	TCGAATCATCAACCAACCTCCGTCCTCCCAACTTGTCTCTGTGTAATCGCTGGAATGTG	720
Db		672	TCGAATCATCAACCAACCTCTTGCTCTCCCAACTTGTCTCTGTGTAATCGCTGGAATGTG	731
OY		721	CGGCGTTTATCATATCTTCCTCTTCAACCTGCGCTATGACCTCATCTTCTTGTTGGTCTT	780
Db		732	CGGCGTTTATCATATCTTCTCTTCAATCTGCTGCTAATGCTCATCTTCTTGTTGGTCTT	791
OY		781	CTGACATCAAGGATATGTGCCGCTTGTGTCTCTAATTCCAGATCTTCAACCAACGACG	840
Db		792	CTGACATCAAGGATATGTGCCGCTTGTGTCTCTCTAATTCCAGATCTTCAACCAACGACG	851
OY		841	AOGGACCATGACGAGCCCTGACGACTCCTGTCAAGGAACCTCATATGTAATCCCTCCGT	900
Db		852	AOGGACCATGCGGAGCCCTGACGACTCTGTCAAGGAACCTCATATGTAATCCCTCCGT	911
OY		901	TGCTGTACAAAACTTCGAGTAGAAGCTGACCTGTATTCGCATCCCATCATCTCGGACT	960
Db		912	TGCTGTACAAAACTTCGAGTAGAAGCTGTATTCGCATCCCATCATCTCGGACT	971
OY		961	TTGCGAAAAATCTATAGGAGTGGGCGCTCAGCCGCTTCTCTCGGCTCAGTTTACTAGTG	1020
Db		972	TTGCGAAAAATCTATAGGAGTGGGCGCTCAGCCGCTTCTCTCGGCTCAGTTTACTAGTG	1031
OY		1021	CCATTGTTCACTGGTGTGTAAGGCGTTTCCCCACACTGTATGGCTTTCAGTATATANGATG	1080
Db		1032	CCATTGTTCACTGGTGTGTAAGGCGTTTCCCCACACTGTATGGCTTTCAGTATATANGATG	1091
OY		1081	ATGTTGTACTGGGGGCCAAGTCTG-TACACCATCTTGAGTCCCTTTTACCCTGTATACC	1139
Db		1092	ATGTTGTACTGGGGGCCAAGTCTGTGTACAGATCTTGAGTCCCTTTTACCCTGTATACC	1151
OY		1140	AATTTCCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA	1181
Db		1152	AATTTCCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA	1193
RESULT 10				
AB074844				
LOCUS		1170 bp	DNA	linear VRL 28-MAY-2002
DEFINITION		Hepatitis B virus gene for large S protein, complete cds,		
ACCESSION		AB074844.		
VERSION		AB074844.1	GI:21218033	
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS		1	Kato,H., Orito,E., Suganuchi,F., Ueda,R., Koshizaka,T., Yanaka,S., Gish,R.G. and Mizokami,M.	
TITLE			New hepatitis B virus genotyping system that allows for identification of mixed genotype infection	
JOURNAL			Unpublished	
REFERENCE		2	(bases 1 to 1170)	
AUTHORS			Hideaki,K.	
TITLE			Direct Submission	
JOURNAL			Submitted (25-NOV-2001) Kato Hideaki, Nagoya City University, Second Department of Medicine, Kawasumi 1, Mizuhō-cho, mizuho-ku, Nagoya, Aichi 467-0001, Japan (E-mail:hideakixnagoya2.jrc.or.jp, Tel:81-52-853-8216)	
FEATURES			Location/Qualifiers	
SOURCE			1..1170	/organism="Hepatitis B virus"

	CDS	/mol type="genomic DNA"	/db xref="taxon:10407"	/clone="BNG48"	1..1170
					/codon_start=1
					/product="large S protein"
					/protein_id="BAB96531.1"
					/db_xref="GI:21218034"
					/translation="MGONLSTSNPLGFPPHOLDPARAANTAHEDDNPENKOTWPDPA NKGAVAGFGLTPEPHGDLGWSPQAGIILQLTPANPAPASTNSOGSOPLPSPLE NLTHQAQRNSTFHOTLODPVRVGLIPPGSGSSGTVPNTIVSHISLSIGDE ALNNETITSGFLGPLVLIOAGFPFLTRILRIPOSIDMWTSLNFAGTTVLGONSOS PTSHSPSCPCPTCPGRYMMCLRRFIIFLFLLICLIFFLVLDYOGLMPCPPIPGS STSHTSCRTCTNPABGTSMYPGCCCKPKPDGNCTCIPISSMAFGRLMEASARPS" MSSLIVAFVQMFTGVSVSEPTWLISVIAMMWGPISLYLSLPPILPIPFCLMYTI"
BASE COUNT	249 a	345 c	254 g	322 t	
ORIGIN					
Query Match	Best Local Similarity	94.5%	Score 1115.6;	DB 14;	Length 1170;
Matches 1136;	Conservative	0;	Mismatches 34;	Indels 0;	Gaps 0;
Dy	1	ATGGGGAGAAGATCTTTCCACCGACGAATCCTTGCGATTCTTCCCAGCACCAAGTTGGAT	60		
Dd	1	ATGGGGAGAAGATCTTTCCACCGACGAATCCTTGCGATTCTTCCCAGCACCAAGTTGGAT	60		
Oy	61	CCAGCTTCAGAGCAAACCACAATCCAGATTGGACTTCCTAAATCCCAAGAACACC	120		
Dd	61	CCAGCTTCAGAGCAAACCACAATCCAGATTGGACTTCCTAAATCCCAAGAACACC	120		
Oy	121	TGGCCAAACCGCCAAGAATAAGTAGAGCTGGAGCATTCGGAGCTGGGATTACCCCCAGC	180		
Dd	121	TGGCCAAACCGCCAAGAATAAGTAGAGCTGGAGCATTCGGAGCTGGGATTACCCCCAGC	180		
Oy	181	GGAGAGCTTTTGGGGTGAAGCCCTCAGAGCTCAGGGCAATTAACAACAACCTTGCAGCAAT	240		
Dd	181	GGAGAGCTTTTGGGGTGAAGCCCTCAGAGCTCAGGGCAATTAACAACAACCTTGCAGCAAT	240		
Oy	241	CCGCTCTCTGTTCCACCAATCGCAGTCAAGGAAGGACCTACCCGCTGTCTCAACT	300		
Dd	241	CCGCTCTCTGTTCCACCAATCGCAGTCAAGGAAGGACCTACCCGCTGTCTCAACT	300		
Oy	301	TTGAGAAAACACTCATCTCAAGCCATGCAATGGAATTCACACATTTTCACCAAACTCTG	360		
Dd	301	TTGAGAAAACACTCATCTCAAGCCATGCAATGGAATTCACACATTTTCACCAAACTCTG	360		
Oy	361	CAAAATCCGAGAGGAGAGTCTGTATTTCCGTCGTGTGCTCAGTTCAAGAAACAGTA	420		
Dd	361	CAAAATCCGAGAGGAGAGTCTGTATTTCCGTCGTGTGCTCAGTTCAAGAAACAGTA	420		
Oy	421	AACCTTTCGCACTACTGTCTCTCCCATATCTGCAATCTTCTGAGGATTTGGGAGCCCT	480		
Dd	421	AACCTTTCGCACTACTGTCTCTCCCATATCTGCAATCTTCTGAGGATTTGGGAGCCCT	480		
Oy	481	GCGCGAAACATGGAACATCAATCAGAAATTCCTAGAGCCCTCTGTGTTCACGGCG	540		
Dd	481	GCGCGAAACATGGAACATCAATCAGAAATTCCTAGAGCCCTCTGTGTTCACGGCG	540		
Oy	541	GGGTTTTTTCTGTGTGAACAAGATCTTCAAAATACCGCAGATCTNAATCTGTGTGACT	600		
Dd	541	GGGTTTTTTCTGTGTGAACAAGATCTTCAAAATACCGCAGATCTNAATCTGTGTGACT	600		
Oy	601	TCTCTCAATTTTCTAGAGGGGAACCTACCGTGTCTTTGAGCCAAAAATTTGGAGTCCCAAC	660		
Dd	601	TCTCTCAATTTTCTAGAGGGGAACCTACCGTGTCTTTGAGCCAAAAATTTGGAGTCCCAAC	660		
Oy	661	TCGAATCACTACCAACCTCTGTCTTCAACTTTGCTGGTTATTCGCTGATGTGTCTG	720		
Dd	661	TCGAATCACTACCAACCTCTGTCTTCAACTTTGCTGGTTATTCGCTGATGTGTCTG	720		
Oy	721	CGGGGTTTATCAATCTTCTCTTCAATCTCTGTGTGTATAGCTCAATCTTCTGTGTGTCTT	780		
Dd	721	CGGGGTTTATCAATCTTCTCTTCAATCTCTGTGTGTATAGCTCAATCTTCTGTGTGTCTT	780		



QY 961 TTCGAAAATTCCTATGGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTACTAGT 1020  
DB 961 TTCGAAAATTCCTATGGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTACTAGT 1020  
QY 1021 CCATTGTTCAAGTGGTTCGTAGAGCTTTTCCCCACGTTTGGCTTCACTATATGATG 1080  
DB 1021 CCATTGTTCAAGTGGTTCGTAGAGCTTTTCCCCACGTTTGGCTTCACTATATGATG 1080  
QY 1081 ATGTGTACTGGGGGCAAGTCTGTACACATCTGAGTCCCTTTTACCGCTTACCA 1140  
DB 1081 ATGTGTATGGGGGCAAGTCTGTACACACCTTGAAGTCCCTTTTACCGCTTACCA 1140  
QY 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170  
DB 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170

RESULT 12  
AB074842 1170 bp DNA linear VRL 28-May-2002  
LOCUS Hepatitis B virus gene for large S protein, complete cds,  
DEFINITION clone:ENG08.  
ACCESSION AB074842  
VERSION AB074842  
KEYWORDS  
SOURCE Hepatitis B virus  
ORGANISM Hepatitis B virus  
REFERENCE  
AUTHORS Kato,H., Orito,E., Suganuchi,F., Ueda,R., Koshizaka,T., Yanaka,S.,  
Gish,R.G. and Mizokami,M.  
TITLE New hepatitis B virus genotyping system that allows for  
identification of mixed genotype infection  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1170)  
AUTHORS Hideaki,K.  
TITLE Direct Submission  
JOURNAL Submitted (25-NOV-2001) Kato Hideaki, Nagoya City University,  
Second Department of Medicine, Kawasaki 1, Mizuho-cho, mizuho-ku,  
Nagoya, Aichi 467-0001, Japan (E-mail:hideaki@nagoya2.jtc.or.jp,  
Tel:81-52-853-8216)

FEATURES  
source  
1..1170  
/organism="Hepatitis B virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10407"  
/clone="BNG08"  
1..1170  
/codon\_start=1  
/product="large S protein"  
/protein\_id="BAB96529.1"  
/db\_xref="GI:21218030"  
/translation="MGOMLSTNPLGFPPDQHDLPARANTANPMDNPNKDTWPA  
NKVGAGAGLFTPHGGLGRSPOAQGLIQLPANPASTNROSGRPTLSPLR  
DHPGAMQNSFTFHTLODPGRVGYPPAGSSGQVNPATTVSHLSSTLRSP  
ARNMNTSGFLPILVLAQGFLLTRILITPOSIDSMWTSINPGLTVCQNSOS  
PTSNSTPSCPTCGRYRMCRLRIIFILLICILFLVLDYQMLPVCPLPFGS  
STTSGCKCTTPPAQGTSMYSCCTKPSDNCCTCIPSPSWAFGKELWEVASRFS  
WLSLVPVQWVGSLPIVWLSIVIMMMWGPSLVSILNPLPLPIFFCLMVYI"

CDS  
247 a 349 c 255 g 319 t

BASE COUNT 247 a 349 c 255 g 319 t  
ORIGIN

Query Match 94.2%; Score 1112.4; DB 14; Length 1170;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1134; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 ATGGGACAGATCTTCCACGAGCAATCTCTGGGATCTTTCCGACACAGTTGAT 60  
DB 1 ATGGGACAGATCTTCCACGAGCAATCTCTGGGATCTTTCCGACACAGTTGAT 60  
QY 61 CCAGCTTCAGAGCAACCAACATCCAGATTGGACTTCATCCACAGACACC 120  
DB 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170

DB 61 CCAGCTTCAGAGCAACCAACATCCAGATTGGACTTCATCCACAGACACC 120  
QY 121 TGGCCAGACGCGCAACAGGTAGAGCTGAGATTCGAGCTGGGGTTACCCCAACGAC 180  
DB 121 TGGCCAGACGCGCAACAGGTAGAGCTGAGATTCGAGCTGGGGTTACCCCAACGAC 180  
QY 181 GGAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACCTTGGCAGCAAT 240  
DB 181 GGAGGCTTTTGGGGGAGGCCCTCAGGCTCAGGGCATTAACAACCTTGGCAGCAAT 240  
QY 241 CCGCTCTGCTTCCACCAATGCGCAGTCAAGAGGCAAGCTTACCCTGTCTCCACT 300  
DB 241 CCGCTCTGCTTCCACCAATGCGCAGTCAAGAGGCAAGCTTACCCTGTCTCCACT 300  
QY 301 TTGAGAAACATCATCTCAAGCCAGTGGAGTGAAGTCCAGAACTTCCACCAACTCG 360  
DB 301 TTGAGAGACTCATCTCAAGCCAGTGGAGTGAAGTCCAGAACTTCCACCAACTCG 360  
QY 361 CAAGATCCCAAGTGAAGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAAGACAGTA 420  
DB 361 CAAGATCCCGAGTGAAGGCTGTATTTCCCTGCTGGTGGCTCCAGTTCAAGACAGTA 420  
QY 421 AACCTGTTCGACTACTGTCTCTCCCATATGTCATCTTCCAGATTGGGACCT 480  
DB 421 AACCTGTTCGACTACTGTCTCTCCCATATGTCATCTTCCAGATTGGGACCT 480  
QY 481 GGGCGGAACATGGAACATCATCAGAGTTCCTAGAGCCCTGCTGTTACAGGG 540  
DB 481 GGGCGGAACATGGAACATCATCAGAGTTCCTAGAGCCCTGCTGTTACAGGG 540  
QY 541 GGGTTTTCTTGTGACAAAGATTCCTCAATACCGCAGAGTCTAGACTCTGTGTGACT 600  
DB 541 GGGTTTTCTTGTGACAAAGATTCCTCAATACCGCAGAGTCTAGACTCTGTGTGACT 600  
QY 601 TCTCTCAATTTCTAGGGGGAACCTACCGTGTCTTTGGCAAAATTCGACGCCAAC 660  
DB 601 TCTCTCAATTTCTAGGGGGAACCGTGTCTTTGGCAAAATTCGACGCCAAC 660  
QY 661 TCCATTCATCTCAACCACTCTGCTCCCAATGTCCTGTTATCGCGTAGTGTCTG 720  
DB 661 TCCATTCATCTCAACCACTCTGCTCCCAATGTCCTGTTATCGCGTAGTGTCTG 720  
QY 721 CGGCTTTTATCATCTTCTCTTCAATCCGCTGCTATGCTCATCTTCTGTGGTCTT 780  
DB 721 CGGCTTTTATCATCTTCTCTTCAATCCGCTGCTATGCTCATCTTCTGTGGTCTT 780  
QY 781 CTGACATTCAGAGTATGTGCGCTTGTCTCTAATTCAGAGATCTTCAACACACAC 840  
DB 781 CTGACATTCAGAGTATGTGCGCTTGTCTCTAATTCAGAGATCTTCAACACACAC 840  
QY 841 ACGGACCATGAGAGCTTGCACACTCTGCTCAAGGAACCTCTATGATTCCTCTCT 900  
DB 841 ACGGACCATGAGAGCTTGCACACTCTGCTCAAGGAACCTCTATGATTCCTCTCT 900  
QY 901 TGCCTGACAAACCTTCGATGGAATGCACTGATATTCATCCATCATCTTGGGCT 960  
DB 901 TGCCTGACAAACCTTCGAGGGAATGCACTGATATTCATCCATCATCTTGGGCT 960  
QY 961 TTCGAAAATTCCTATGGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTACTAGT 1020  
DB 961 TTCGAAAATTCCTATGGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTACTAGT 1020  
QY 1021 CCATTGTTCAAGTGGTTCGTAGAGCTTTTCCCCACGTTTGGCTTCACTATATGATG 1080  
DB 1021 CCATTGTTCAAGTGGTTCGTAGAGCTTTTCCCCACGTTTGGCTTCACTATATGATG 1080  
QY 1081 ATGTGTACTGGGGGCAAGTCTGTACACACATCTGAGTCCCTTTTACCGCTTACCA 1140  
DB 1081 ATGTGTATGGGGGCAAGTCTGTACACACATCTTGAATCCCTTTTACCGCTTACCA 1140  
QY 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170  
DB 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170

RESULT 13  
 HBUS5226  
 LOCUS HBUS5226 1149 bp DNA linear VRL 15-FEB-1997  
 DEFINITION Hepatitis B virus large S protein (Pre-S/S) gene, partial cds.  
 ACCESSION U55226  
 VERSION U55226.1 GI:1518558  
 KEYWORDS  
 SOURCE Hepatitis B virus  
 ORGANISM Hepatitis B virus  
 REFERENCE 1 (bases 1 to 1149)  
 AUTHORS Moraes,M.T., Gomes,S.A. and Niel,C.  
 TITLE Sequence analysis of pre-S/S gene of hepatitis B virus strains of genotypes A, D, and F isolated in Brazil  
 JOURNAL Arch. Virol. 141 (9), 1767-1773 (1996)  
 MEDLINE 97049068  
 PUBMED 8893798  
 REFERENCE 2 (bases 1 to 1149)  
 AUTHORS Niel,C.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-APR-1996) Christian Niel, Virology, Oswaldo Cruz Foundation, Av. Brasil, 4365, Rio de Janeiro, RJ 21045-900, Brazil  
 FEATURES  
 source  
 1. 1149  
 /organism="Hepatitis B virus"  
 /mol\_type="genomic DNA"  
 /strain="BrM (ayw2)"  
 /db\_xref="taxon:10407"  
 1. 1149  
 /gene="Pre-S/S"  
 1. >1149  
 /gene="Pre-S/S"  
 /note="surface antigen"  
 /product="large S protein"  
 /protein\_id="AAB47472.1"  
 /db\_xref="GI:1518559"  
 /translation="MGONLSTNPLGFPPHOLDPAFRANTANDMDFNPDTPDA  
 NKYGAFLGIGTPPHGGLGNSPQAGIITQTPANPPASTNRQROTPSPR  
 NHPQAMQNSTPHOTLQDPRVGLTFPAGSSSGTVNVPPTASTISIFRIGDP  
 ALNMENITSGFLPVLQAGFLTLTILTIPOSLSWMSLNFLLGTTVCLONSGS  
 PLSNHSPTSCPTPCPGRYMMLRFLIFLILLCILFLVLVDYQGLMPLVCPILPS  
 STSTGCRCTCTPAQCTSMYPSGCCCTKPSDGNCTCIPSSNAFGFLMEMASRPS  
 WSLVDPVQMFGLSPTVWLVIYIMMMWVGPISLISLFLPLPIFF"  
 BASE COUNT 240 a 350 c 249 g 310 t  
 ORIGIN  
 Query Match 94.0%; Score 1110.6; DB 14; Length 1149;  
 Best Local Similarity 97.9%; Pred. No. 0;  
 Matches 1125; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 1 ATGGGCGAAGATCTTTCCACGACGATCTCTGGGATTTTCCCGACACCACTTGAT 60  
 DB 1 ATGGGCGAAGATCTTTCCACGACGATCTCTGGGATTTTCCCGACACCACTTGAT 60  
 QY 61 CCAAGCTTCAGAGCAACCAACATCCAGATTGGGACTTATCCCAACAGGACACC 120  
 DB 61 CCAAGCTTCAGAGCAACCAACATCCAGATTGGGACTTATCCCAACAGGACACC 120  
 QY 121 TGGCCAGACGCAACAGTAGAGCTGAGCATTCGGAATGGGGTTACCCCAACCGCAC 180  
 DB 121 TGGCCAGACGCAACAGTAGAGCTGAGCATTCGGAATGGGGTTACCCCAACCGCAC 180  
 QY 181 GGAAGCCTTTTGGGGTGAAGCCTCAGGCTCAGGGCATTAACAACAACCTTGCCGCAAT 240  
 DB 181 GGAAGCCTTTTGGGGTGAAGCCTCAGGCTCAGGGCATTAACAACAACCTTGCCGCAAT 240  
 QY 241 CCGCTCTGCTTCCACCAATGCGCAGTCAGGAGGAGAGCCTTACCCGCTGTCTCCACT 300  
 DB 241 CCGCTCTGCTTCCACCAATGCGCAGTCAGGAGGAGAGCCTTACCCGCTGTCTCCACT 300

QY 301 TTGAGAAACATCATCTCCAGGCCATGAGTGAACTCCACAACCTTCCACAACACTCTG 360  
 DB 301 TTGAGAAACATCATCTCCAGGCCATGAGTGAACTCCACAACCTTCCACAACACTCTG 360  
 QY 361 CAAGATCCCAAGATGAGAGTCTGTATTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 DB 361 CAAGATCCCAAGATGAGAGTCTGTATTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 QY 421 AACCTGTTCCGACTAGCTCTCTCCCATATGCTCAATCTTTCAGAGATTGGGAGACCT 480  
 DB 421 AACCTGTTCCGACTAGCTCTCTCCCATATGCTCAATCTTTCAGAGATTGGGAGACCT 480  
 QY 481 GCGCGGAACATGGAACAATCATCATGAGATTCTAGAACCCCTGCTGCTGCTGCTGCTG 540  
 DB 481 GCGCGGAACATGGAACAATCATCATGAGATTCTAGAACCCCTGCTGCTGCTGCTGCTG 540  
 QY 541 GGGTTTTTCTGTTGACCAAGATTCCTCAAAATCCCAAGAGTTTAACTCGTGTGACT 600  
 DB 541 GGGTTTTTCTGTTGACCAAGATTCCTCAAAATCCCAAGAGTTTAACTCGTGTGACT 600  
 QY 601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 660  
 DB 601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 660  
 QY 661 TCCAACTACTCACCAACTCTCTGCTCCCAACTGTCTGTATGCTGATGCTGCTG 720  
 DB 661 TCCAACTACTCACCAACTCTCTGCTCCCAACTGTCTGTATGCTGATGCTGCTG 720  
 QY 721 CCGCGTTTATCATCTTCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 DB 721 CCGCGTTTATCATCTTCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 781 CTGAGATCAAGATGATGTTGCGGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 DB 781 CTGAGATCAAGATGATGTTGCGGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY 841 ACGGAGACATGACAGCTGACAGCTCTGCTCAAGAAACCTGTATGATCCCTCTGT 900  
 DB 841 ACGGAGACATGACAGCTGACAGCTCTGCTCAAGAAACCTGTATGATCCCTCTGT 900  
 QY 901 TGTCTGACAAACCTTCCGATGGAACCTGACCTGTATCCCATCATCATCTGAGCT 960  
 DB 901 TGTCTGACAAACCTTCCGATGGAACCTGACCTGTATCCCATCATCATCTGAGCT 960  
 QY 961 TTGCGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGAGCTCAGTTACTAGT 1020  
 DB 961 TTGCGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGAGCTCAGTTACTAGT 1020  
 QY 1021 CCAATTGTTCAAGTGTTCGATGAGGCTTTCCCACTGTTTGGCTTCAAGTTATATGATG 1080  
 DB 1021 CCAATTGTTCAAGTGTTCGATGAGGCTTTCCCACTGTTTGGCTTCAAGTTATATGATG 1080  
 QY 1081 ATGTTGACTGGGGGCCAAGTCTGTACACATCTTGAGTCCCTTTTAAACCGCTTAACA 1140  
 DB 1081 ATGTTGACTGGGGGCCAAGTCTGTACACATCTTGAGTCCCTTTTAAACCGCTTAACA 1140  
 QY 1141 ATTTTCTTT 1149  
 DB 1141 ATTTTCTTT 1149  
 RESULT 14  
 HBUS5225  
 LOCUS HBUS5225 1149 bp DNA linear VRL 15-FEB-1997  
 DEFINITION Hepatitis B virus large S protein (Pre-S/S) gene, partial cds.  
 ACCESSION U55225  
 VERSION U55225.1 GI:1518556  
 KEYWORDS  
 SOURCE Hepatitis B virus  
 ORGANISM Hepatitis B virus  
 REFERENCE 1 (bases 1 to 1149)  
 AUTHORS Moraes,M.T., Gomes,S.A. and Niel,C.

TITLE Sequence analysis of pre-S/S gene of hepatitis B virus strains of  
JOURNAL genotypes A, D, and F isolated in Brazil  
MEDLINE Arch. Virol. 141 (9), 1767-1773 (1996)  
PUBMED 97049068  
8893798  
REFERENCE 2 (bases 1 to 1149)  
AUTHORS Nêl, C.  
JOURNAL Direct Submission  
Submitted (17-APR-1996) Christian Nêl, Virology, Oswaldo Cruz  
Foundation, Av. Brasil, 4365, Rio de Janeiro, RJ 21045-900, Brazil

FEATURES  
source  
1..1149  
/organism="Hepatitis B virus"  
/mol\_type="genomic DNA"  
/strain="Brn (ayw2)"  
/db\_xref="taxon:10407"  
1..1149  
/gene="Pre-S/S"  
1..>1149  
/gene="Pre-S/S"  
/note="surface antigen"  
/codon\_start=1  
/product="large S protein"  
/protein\_id="AAB47471.1"  
/db\_xref="GI:1518557"  
/translation="MGONLSTSNPLGFPPDHLDPAPFRANTANPMDENFNDOTWPD  
NKVAGAFGLFTPHGGLGWSPOAGLLOTLPANPASPNTNROSGRPPTLSPLR  
NTHPOMONSTPHOTLODPVRGLXPAGSSSGSTVNPVTTASRSSISIRIGDP  
ALMENTISGRFLPLVLOAGFLRLRIITFQSDSWFTSNFAGTIVCLGQNSQS  
PTSNHSPTSCPTCPGPRMMLRFLIFLILLCLIFLLVLDYQMLPVCPLPGS  
STSTGPCTCTTPAGCTSMPSCCCTKPSDNCCTCIPSPSMWARKKFLMEWASARFS  
WLSLVPFQWTVGLSPYVWLSVIMMMWMPGSLVSLSPFLPLPIPF"

BASE COUNT 239 a 350 c 250 g 310 t

ORIGIN

Query Match 93.8%; Score 1107.4; DB 14; Length 1149;  
Best Local Similarity 97.7%; Pred. No. 0;  
Matches 1123; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGGGGGGAATCTTTCACAGCAATCTCTGGGATTTCTTCCGACACCAAGTGGAT 60  
DB 1 ATGGGGGGAATCTTTCACAGCAATCTCTGGGATTTCTTCCGACACCAAGTGGAT 60  
QY 61 CCAGCCTTCAGAGCAAAACCAACAATCCAGATTGGGACTTCAATCCCAAGAGCACC 120  
DB 61 CCAGCCTTCAGAGCAAAACCAACAATCCAGATTGGGACTTCAATCCCAAGAGCACC 120  
QY 121 TGGCCAGAGCGCAACAGGTAGAGCTGAGCAATTCGAGCTGGGGTTCAACCCGAC 180  
DB 121 TGGCCAGAGCGCAACAGGTAGAGCTGAGCAATTCGAGCTGGGGTTCAACCCGAC 180  
QY 181 GGAAGGCTTTGGGGTGGAGCCCTCAGGCTCAGGGATTAACAACAACCTTGGCAGCAAT 240  
DB 181 GGAAGGCTTTGGGGTGGAGCCCTCAGGCTCAGGGATTAACAACAACCTTGGCAGCAAT 240  
QY 241 CGGCTCTCTGCTTCACCAATCGCCAGTGAAGAGGACAGTCAACCCGGTGTCCACCT 300  
DB 241 CGGCTCTCTGCTTCACCAATCGCCAGTGAAGAGGACAGTCAACCCGGTGTCCACCT 300  
QY 301 TTGAGAAACACTATCTTCAGCCATGAGTGAAGTTCACCACTTTCACCAAACTCTG 360  
DB 301 TTGAGAAACACTATCTTCAGCCATGAGTGAAGTTCACCACTTTCACCAAACTCTG 360  
QY 361 CAAGATCCCAAGAGTGAAGAGTGTATTTTCTGCTGGTGGCTCCAGTTTCAAGAACAT 420  
DB 361 CAAGATCCCAAGAGTGAAGAGTGTATTTTCTGCTGGTGGCTCCAGTTTCAAGAACAT 420  
QY 421 AACCGTTCGAGTACTGTCTCTCCCATATGTCATTTCTCGAGGATGGGGACCT 480  
DB 421 AACCGTTCGAGTACTGTCTCTCCCATATGTCATTTCTCGAGGATGGGGACCT 480  
QY 481 GGGCGAATGAGAGCAATCAATCAGATTCTAGAACCCCTGCTCGTTTACAGGCG 540

DB 481 GGGCGAATGAGAGCAATCAATCAGATTCTAGAACCCCTTCTCGTTTACAGGCG 540  
QY 541 GGGTTTTCTTGTGACAAAGATCTCAATACCGAGAGCTGAGCTGGTGGACT 600  
DB 541 GGGTTTTCTTGTGACAAAGATCTCAATACCGAGAGCTGAGCTGGTGGACT 600  
QY 601 TCTCTCAATTTCTAGGGGGAACCTACCGTGTCTGAGCAAAATTCGAGTCCCAAC 660  
DB 601 TCTCTCAATTTCTAGGGGGAACCTACCGTGTCTGAGCAAAATTCGAGTCCCAAC 660  
QY 661 TCCAACTACTACCAACCTCTGCTCTCAATTTGCTGTTATGCTGATGTCTG 720  
DB 661 TCCAACTACTACCAACCTCTGCTCTCAATTTGCTGTTATGCTGATGTCTG 720  
QY 721 CGGCGTTTATCATCTCTCTTCAATCCGCTGATAGCTCATCTCTGTGGTCTT 780  
DB 721 CGGCGTTTATCATCTCTCTTCAATCCGCTGATAGCTCATCTCTGTGGTCTT 780  
QY 781 CTGACTATCAAGTATGTTGCGCTTGTCTCTTAATTCAGAGATCTCAACACAC 840  
DB 781 CTGACTATCAAGTATGTTGCGCTTGTCTCTTAATTCAGAGATCTCAACACAC 840  
QY 841 ACGGACATGACAGACCTGACAGACTCTGCTCAAGAACTCTATGATCTCTCT 900  
DB 841 ACGGACATGACAGACCTGACAGACTCTGCTCAAGAACTCTATGATCTCTCT 900  
QY 901 TCCCTTACAAACCTTCGATGGAACGACCTGATTTCCATCCATCACTCTGGCT 960  
DB 901 TCCCTTACAAACCTTCGATGGAACGACCTGATTTCCATCCATCACTCTGGCT 960  
QY 961 TTCGAAATTCCTATGGAAGTGGGCTCAGCCCTTCTCTGCTCAGTTACTAGT 1020  
DB 961 TTCGAAATTCCTATGGAAGTGGGCTCAGCCCTTCTCTGCTCAGTTACTAGT 1020  
QY 1021 CCATTTGTCAGTGTCTGATAGGCTTTTCCCACTGTTTGGCTTCAATATGATG 1080  
DB 1021 CCATTTGTCAGTGTCTGATAGGCTTTTCCCACTGTTTGGCTTCAATATGATG 1080  
QY 1081 ATGTTGTAATGAGGAGCAAGTGTGACACATCTGATGCTTTTACCGTGTACA 1140  
DB 1081 ATGTTGTAATGAGGAGCAAGTGTGACACATCTGATGCTTTTACCGTGTACA 1140  
QY 1141 ATTTTCTT 1149  
DB 1141 ATTTTCTT 1149

RESULT 15  
HBVPS12SP 1401 bp DNA linear VRL 01-FEB-1994  
LOCUS  
DEFINITION Hepatitis B virus Pre S1, Pre S2 and S genes (patient P).  
ACCESSION X77310.1 GI:452615  
VERSION S gene; S protein; S1 gene; S1 protein; S2 gene; S2 protein;  
KEYWORDS surface protein.  
SOURCE  
ORGANISM Hepatitis B virus  
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
REFERENCE  
1.  
AUTHORS Lai, M.E., Nazzari, A.P. and Balistreri, A.  
TITLE Sequence analysis of HBV genomes isolated from patients with HbsAg  
JOURNAL Unpublished  
2 (bases 1 to 1401)  
REFERENCE  
Lai, M.E.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (19-JAN-1994) M.E. Lai, University of Cagliari, Institute  
of Internal Medicine, Laboratory of Virology, Via San Giorgio 12,  
09124 Cagliari, ITALY  
LOCATION/Qualifiers  
1..1401  
/organism="Hepatitis B virus"  
/mol\_type="genomic DNA"



```

/sub strain="ayw"
/isolate="patient p"
/db xref="taxon:10407"
/clone="3 overlapping PCR clones"
38..1207
/gene="pres1/pres2 and S orf"
38..1207
/gene="pres1/pres2 and S orf"
/codon_start=1
/product="HBV surface proteins"
/protein_id="CAA54516.1"
/db_xref="GI:452616"
/db_xref="SPTREMBL:Q67956"
/translation="MCQNLSTSNPLGFPPHOLDPAFRANTANPDMPNPKDTPDA
NKYGAGFGFTFPHGLGMSPOAGLIQTVANPPASTNOSGROPTPLR
NTHPOMOMSTPHOTLDPVRGLYFPRVNGSSGTVNPTTVSHISSIPARTGP
VTMENTSGFLPQVLYQNGPFLITLITPOSIDBWMVSNFLGTTTCLONSBS
PTSNHSPSCPTCPGIRKMSRRFIIIFLILCLIFLLVDLYQGLMFLVCPILRS
STSTGECRTCTTPAOGTSMYPSCCCTKPDGNCCTIPISVNAFGKFLMWSARPS
WLSLVGFVQMFVGLSPTVWLGIIMMMYMGHSLYSLISPLPLPIFCLMVIYI"
BASE COUNT      316 a      385 c      307 g      393 t
ORIGIN
Query Match      93.8%; Score 1107.4; DB 14; Length 1401;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1135; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY      1 ATGGGGGAGATCTTTCCACAGCAATCTGCGATTTCTCCGACCAACAGTGGAT      60
DB      38 ATGGGGGAGATCTTTCCACAGCAATCTGCGATTTCTCCGACCAACAGTGGAT      97
QY      61 CCAAGCTTCAGAGCAACACCAACATCCAGATTGGGACTTCAATCCCAACAGAGACC      120
DB      98 CCAAGCTTCAGAGCAACACCGCAATCCAGATTGGGACTTCAATCCCAACAGAGACC      157
QY      121 TGGCCAGACGCGCAACAAAGTAGAGCTGAGCAATTCGAGCTGGGGTTCACCCAACGGC      180
DB      158 TGGCCAGACGCGCAACAAAGTAGAGCTGAGCAATTCGAGCTGGGGTTCACCCAACGGC      217
QY      181 GGAAGCGCTTTGGGGTGAAGCCCTCAGGCTCAGGAGCAATACACAAACCTTGGCAGCAAT      240
DB      218 GGAAGCGCTTTGGGGTGAAGCCCTCAGGCTCAGGAGCAATACACAAACCTTGGCAGCAAT      277
QY      241 CCGGCTCTGCTTCCACCAATCGCCAGTCAAGAGGAGCACTTACCCGCTGTCTCAACT      300
DB      278 CCGGCTCTGCTTCCACCAATCGCCAGTCAAGAGGAGCACTTACCCGCTGTCTCAACT      337
QY      301 TTGAGAAACACTCTCTCAAGCATGCAAGTGAATTCACAACTTTCCACCAACTCTG      360
DB      338 TTGAGAAACACTCTCTCAAGCATGCAAGTGAATTCACAACTTTCCACCAACTCTG      397
QY      361 CAAGATCCAGAGTGAAGGCTGTATTTCCCTGCTGGTGCACAGTTCCAGAAACAGTA      420
DB      398 CAAGATCCAGAGTGAAGGCTGTATTTCCCTGCTGGTGCACAGTTCCAGAAACAGTA      457
QY      421 AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTCTGAGGATTGGGAGCCT      480
DB      458 AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTCTGAGGAGTGGGAGCCT      517
QY      481 GCGGGAACATGAGGAACATCAATCAGGATTCTTAGGACCCCTGCTGTGTACAGCG      540
DB      518 GTCAGCATCAGGAACATCAATCAGGATTCTTAGGACCCAGCTGTGTGTACAGCG      577
QY      541 GGGTTTTTCTGTGACAAAGATCTCAATACCGAGAGTCTAGACTGCTGGTGGACT      600
DB      578 GGGTTTTTCTGTGTGACAAAGATCTCAATACCGAGAGTCTAGACTGCTGGTGGACT      637
QY      601 TCTCTCAATTTTCTAGGGGAACTACCGTGTGTCTTGGCCAAATTCGCAATCCCAACC      660
DB      638 TCTCTCAATTTTCTAGGGGAACTACCGTGTGTCTTGGCCAAATTCGCAATCCCAACC      697
QY      661 TCCATCACTACCAACTCTGTGTCTCAACTGTCTGTGTATCGCTGGATGTGTCTG      720

```

```

DB      698 TCCAAATCACTACCAACCTCTCTGTCTCTCAACTTGTCTGTGTATGCTGGATGACTCCT      757
QY      721 CCGCGTTTTATCATCTTCTCTCTCATCTGCTGCTAGTGGCTCATCTTGTGTGTCTT      780
DB      758 CCGCGTTTTATCATCTTCTCTCTCATCTGCTGCTAGTGGCTCATCTTGTGTGTCTT      817
QY      781 CTGAGACTATCAGGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAACGAC      840
DB      818 CTGAGACTATCAGGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAACGAC      877
QY      841 ACGGGAACATGACAGAGCTGCAAGACTCCGCTCAAGGAACCTTATGATCCCTCCTGT      900
DB      878 ACGGGAACATGACAGAGCTGCAAGACTCCGCTCAAGGAACCTTATGATCCCTCCTGT      937
QY      901 TGCTGTACAAACCTTCGATGGAACCTGACCTGTATTCATCCCATCATCTGAGGCT      960
DB      938 TGCTGTACAAACCTTCGATGGAACCTGACCTGTATTCATCCCATCATCTGAGGCT      997
QY      961 TTGGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTACTAGTG      1020
DB      998 TTGGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTACTAGTG      1057
QY      1021 CCAATTTGTCAGTGTTCGTAAGGCTTTCCCACTGTTGGCTTCAGTTATATGAGTG      1080
DB      1058 GCAATTTGTCAGTGTTCGTAAGGCTTTCCCACTGTTGGCTTCAGTTATATGAGTG      1117
QY      1081 ATGTGTACTGGGGGCAAGTCTGTACACCATCTTGAAGTCCCTTTTACCGCTGTACCA      1140
DB      1118 ATGTGTACTGGGGGCAAGTCTGTACACCATCTTGAAGTCCCTTTTACCGCTGTACCA      1177
QY      1141 ATTTCTTTTGTCTTGGGTATACATTTAAACCTTAATAA      1181
DB      1178 ATTTCTTTTGTCTTGGGTATACATTTAAACCTTAATAA      1218

```

Search completed: September 15, 2003, 02:24:48  
 Job time : 3088 secs





XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Coleman PF, Mushahwar IA;  
 XX  
 XX WPI; 2003-040642/03.  
 XX P-PSDB; ABP55073.  
 XX  
 PT New nucleotide sequences, useful for detecting compounds that bind to  
 PT SP41 protein or antibodies to the Hepatitis B virus surface antigen  
 PT mutant or for creating primers and probes .  
 XX  
 PS Claim 1; Fig 1; 43pp; English.  
 XX  
 CC The present sequence is the nucleotide sequence of the entire  
 CC envelope gene for a mutant hepatitis B virus (HBV) strain isolated  
 CC from a French sample identified as 990525169. This is an HBV  
 CC subtype ayw2, genotype D sequence which contains mutations leading  
 CC to 3 amino acid substitutions in the encoded protein: Thr to Ala  
 CC 123, which affects the H166 epitope; Tsp to Leu 199, which is  
 CC outside the 'a' determinant; and Ser to Thr 207, which is also  
 CC outside the 'a' determinant. The present invention relates to the  
 CC novel HBV mutant which has a modified 'a' determinant as a result  
 CC of T123A amino acid substitution, and to methods of detecting this  
 CC mutant, and/or antibodies to the mutant, in patient samples. The  
 CC identification and detection of mutant HBV may lead to improved  
 CC vaccine development and detection systems.  
 CC  
 XX Sequence 1181 BP; 256 A; 351 C; 252 G; 322 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 1181; DB 25; Length 1181;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 TCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660  
 |||||  
 DB 601 TCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660  
 QY 661 TCCAAATCACTCAACAACTCTCTCTCCAACTTGTCTGTGTATGTGGTGGATGTC 720  
 |||||  
 DB 661 TCCAAATCACTCAACAACTCTCTCTCCAACTTGTCTGTGTATGTGGTGGATGTC 720  
 QY 721 CGGCGTTTATCAATCTTCCCTTCATCCGCTGCTAGTGGCTCATCTTCTTGGTCTT 780  
 |||||  
 DB 721 CGGCGTTTATCAATCTTCCCTTCATCCGCTGCTAGTGGCTCATCTTCTTGGTCTT 780  
 QY 781 CTGGACTATCAAGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACACAGC 840  
 |||||  
 DB 781 CTGGACTATCAAGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACACAGC 840  
 QY 841 ACGGACATGACAGAGCTGACAGCTCTGCTCAAGAACTCTATATATCCCTCTGT 900  
 |||||  
 DB 841 ACGGACATGACAGAGCTGACAGCTCTGCTCAAGAACTCTATATATCCCTCTGT 900  
 QY 901 TGGCTATCAAAACCTTCGATGGAATGGAACCTGATTCCTCATATCCTGAGCT 960  
 |||||  
 DB 901 TGGCTATCAAAACCTTCGATGGAATGGAACCTGATTCCTCATATCCTGAGCT 960  
 QY 961 TTGGAAAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCACTTACTAGT 1020  
 |||||  
 DB 961 TTGGAAAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCACTTACTAGT 1020  
 QY 1021 CCATTGTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1080  
 |||||  
 DB 1021 CCATTGTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1080  
 QY 1081 ATGTGTACAGGGGGGCAAGTGTGATCACCATCTGAGTCCCTTTTACCGCTTACCA 1140  
 |||||  
 DB 1081 ATGTGTACAGGGGGGCAAGTGTGATCACCATCTGAGTCCCTTTTACCGCTTACCA 1140  
 QY 1141 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 1181  
 |||||  
 DB 1141 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 1181

RESULT 2  
 AAN00003  
 ID AAN00003 standard; DNA; 2743 BP.  
 XX  
 AC AAN00003;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 14-OCT-1992 (first entry)  
 XX  
 DE Sequence of a part of a hepatitis B virus genome.  
 XX  
 KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine; ss.  
 OS  
 XX Hepatitis B virus.  
 XX  
 FH Key  
 FT CDS 1..84 Location/Qualifiers  
 FT FT /\*tag= a  
 FT FT /label= leader sequence; AAP00006  
 FT FT /note= "reading frame 1"  
 FT FT 88..639  
 FT FT /\*tag= b  
 FT FT /product= AAP00041; core Ag  
 FT FT /note= "reading frame 1"  
 FT FT 1524..2204  
 FT FT /\*tag= c  
 FT FT /product= AAP00042; surface Ag  
 FT FT /note= "reading frame 3"  
 XX  
 XX EP13828-A.  
 XX

PD 06-AUG-1980.  
 XX 21-DEC-1979; 79EP-0303017.  
 XX 22-DEC-1978; 78GB-0049901.  
 XX 22-DEC-1978; 78GB-0049907.  
 PR 27-DEC-1978; 78GB-0050039.  
 PR 01-NOV-1979; 79GB-0037910.  
 XX  
 PA (BIOJ ) BIOGEN NV.  
 PI Murray K, Schaller HE;  
 DR WPI, 1980-57268C/33.  
 DR P-PSDB; AAP00006, AAP00041, AAP00042.  
 XX  
 PT Recombinant DNA coding for polypeptide - have specificity of  
 XX hepatitis B viral antigens in detection or antibody stimulation  
 XX  
 PS Example; Figs 3-9; 43pp; English.  
 XX  
 CC Human serum from a single HBsAg positive, HBeAg positive donor  
 CC (serotype adyn) was used to prep. a DNA-cong. pellet which was  
 CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).  
 CC The labelled DNA was then extracted with phenol from the resulting  
 CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).  
 CC It was then cloned in plasmid pBR322 which was used to transform E.  
 CC coli. Micro-organisms prep'd. by the processes are deposited at the  
 CC NCIB as pBR322-HBv-G-L, e.g. E. coli HB101/pBR322-Pst I dg:  
 CC HBV-kpn I dg: Tetr Amps HBV+.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC  
 SQ Sequence 2743 BP; 638 A; 728 C; 579 G; 798 T; 0 other;

Query Match 95.8%; Score 1131.4; DB 1; Length 2743;  
 Best Local Similarity 97.4%; Pred. No. 0;  
 Matches 1150; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 ATGGGCGAATCTTCCAGCAATCTCTGGGATCTTCCGACCAAGTTGAT 60  
 DB 1035 ATGGGCGAATCTTCCAGCAATCTCTGGGATCTTCCGACCAAGTTGAT 1094  
 QY 61 CCAGCTTCAGAGCAACCAACCAATCCAGTTGGACTTCATCCCAAGACACC 120  
 DB 1095 CCAGCTTCAGAGCAACCAACCAATCCAGTTGGACTTCATCCCAAGACACC 1154  
 QY 121 TGGCCAGACGCCAACAAGTAGAGCTGGAGCATTCGACTGCGGTTCACCCACCGCAC 180  
 DB 1155 TGGCCAGACGCCAACAAGTAGAGCTGGAGCATTCGACTGCGGTTCACCCACCGCAC 1214  
 QY 181 GGAGGCTTTGGGGTGAAGCCCTCAGGCTCAGGGGATACCAAACTTGCAGCAAT 240  
 DB 1215 GGAGGCTTTGGGGTGAAGCCCTCAGGCTCAGGGGATACCAAACTTGCAGCAAT 1274  
 QY 241 CGGCTCTCTGCTTCACCAATGCGCAGTCAAGAGGCAAGCCGCTGTCTCCACT 300  
 DB 1275 CGGCTCTCTGCTTCACCAATGCGCAGTCAAGAGGCAAGCCGCTGTCTCCACT 1334  
 QY 301 TTGAAGAACTCATCTCAGCAAGCATGCACTGCAACTCCAACTTTCACCAAACTCTG 360  
 DB 1335 CTGAAGAACTCATCTCAGCAAGCATGCACTGCAACTCCAACTTTCACCAAACTCTG 1394  
 QY 361 CAAGATCCAGAGTAGAGGTCTGATTTCCGCTGGGGGTCTCCAGTTCAAGAACTAGTA 420  
 DB 1395 CAAGATCCAGAGTAGAGGTCTGATTTCCGCTGGGGGTCTCCAGTTCAAGAACTAGTA 1454  
 QY 421 AACCTGTTCGACTACTGTCTCTCCATATGTCATCTTCTCGAGATTGGGACCTT 480  
 DB 1455 AACCTGTTCGACTACTGTCTCTCCATATGTCATCTTCTCGAGATTGGGACCTT 1514  
 QY 481 GCGCGAATGAGAACTCATCATGAGATTCTAGGACCCCTGCTGTGTTACAGGCG 540  
 DB 1515 GCGCGAATGAGAACTCATCATGAGATTCTAGGACCCCTGCTGTGTTACAGGCG 1574

QY 541 GGGTTTTTCTGTGACAGAAATCTCTCAATACCGCAGAGTCTAGACTCGTGTGACT 600  
 DB 1575 GGGTTTTTCTGTGACAGAAATCTCTCAATACCGCAGAGTCTAGACTCGTGTGACT 1634  
 QY 601 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTGTCTTGGCCAAATTCGACGTCCCAACC 660  
 DB 1635 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTGTCTTGGCCAAATTCGACGTCCCAACC 1694  
 QY 661 TCCATTCATCAACCAACTCCGCTCTCACTGCTCACTGCTGATGCTGATGCTGCTG 720  
 DB 1695 TCCATTCATCAACCAACTCCGCTCTCACTGCTCACTGCTGATGCTGATGCTGCTG 1754  
 QY 721 CGGCTTTTATCATCTTCTCTTCATCTGCTGCTATGCTCATCTTCTGTTGTTCTT 780  
 DB 1755 CGGCTTTTATCATCTTCTCTTCATCTGCTGCTATGCTCATCTTCTGTTGTTCTT 1814  
 QY 781 CTGACTATCAAGATATGTCGCGCTTGTCTCTTAATTCAGAGATCTTCAACCAACGAC 840  
 DB 1815 CTGACTATCAAGATATGTCGCGCTTGTCTCTTAATTCAGAGATCTTCAACCAACGAC 1874  
 QY 841 ACGGACCATGAGAGCCCTGACAGACTGCTGCTCAAGAACTCTATGATCCCTCTGT 900  
 DB 1875 ACGGACCATGAGAGCCCTGACAGACTGCTGCTCAAGAACTCTATGATCCCTCTGT 1934  
 QY 901 TGTCTTACAAAACCTTCGAGATGAACTGCACTGTATTTCCATCCATCATCTGAGGCT 960  
 DB 1935 TGTCTTACAAAACCTTCGAGATGAACTGCACTGTATTTCCATCCATCATCTGAGGCT 1994  
 QY 961 TTGGAAAAATTCATAGGAGTGGGCTCTGAGCCGCTTCTCTGCTGCTCAATTACTAGTG 1020  
 DB 1995 TTGGAAAAATTCATAGGAGTGGGCTCTGAGCCGCTTCTCTGCTGCTCAATTACTAGTG 2054  
 QY 1021 CCATTGTGCTAGTGTGTGATAGGCTTCCCACTGTTGGCTTCACTATATGAGATG 1080  
 DB 2055 CCATTGTGCTAGTGTGTGATAGGCTTCCCACTGTTGGCTTCACTATATGAGATG 2114  
 QY 1081 ATGTGTACTGGGGGCCAAGTCTGTACACCACTTGAATCCCTTTTACCGGTGTAACA 1140  
 DB 2115 ATGTGTACTGGGGGCCAAGTCTGTACACCACTTGAATCCCTTTTACCGGTGTAACA 2174  
 QY 1141 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 1181  
 DB 2175 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 2215

RESULT 3  
 AA004799 standard; DNA; 2743 BP.  
 ID AA004799;  
 AC AA004799;  
 XX  
 AC 25-MAR-2003 (updated)  
 DT 30-OCT-1990 (first entry)  
 XX  
 DE Recombinant sequence encoding hepatitis B antigens.  
 XX  
 KM Hepatitis B virus; vaccine; HBV; ds.  
 XX  
 OS Synthetic.  
 XX  
 OS  
 FH Key Location/Qualifiers  
 FH CDS 1..80  
 FT /\*tag= a  
 FT 87..634  
 FT /\*tag= b  
 FT CDS 1521..2198  
 FT /\*tag= c  
 PN EP374869-A.  
 XX  
 PD 27-JUN-1990.  
 XX



```
FT      /*tag= e
FT      /product= "Polymerase"
FT      1031..2200
FT      /*tag= f
FT      /product= "S1 protein"
FT      1355..2200
FT      /*tag= g
FT      /product= "S2 protein"
FT      1520..2200
FT      /*tag= h
FT      /product= "S protein"
FT      2739..3203
FT      /*tag= i
FT      /product= "X protein"
FT      2955..2965
FT      /*tag= j
FT      /rpl_type= DIRECT
FT      /note= "Direct repeat 2, DR2"
FT      3189..3199
FT      /*tag= k
FT      /rpl_type= DIRECT
FT      /note= "DR1, direct repeat 1"
FT      3212..3272
FT      /*tag= l
FT      /label= 3' epsilon_secondary_structure
FT      3281..3286
FT      /*tag= m
FT      3372..3390
FT      /*tag= n
FT      /note= "SP6 primer binding site"
FT      3391..4089
FT      /*tag= o
FT      /label= Splice_and_polyA_region
FT      5103..5689
FT      /*tag= p
FT      /label= Col_E1_origin
FT      5690..6282
FT      /*tag= q
FT      /label= M13_origin
FT      6462..7405
FT      /*tag= r
FT      /label= Ampicillin_gene
FT      7406..7999
FT      /*tag= s
FT      /label= CMV_promoter
FT      EP1149917-A2.
FT      31-OCT-2001.
FT      20-APR-2001; 2001EP-0303622.
FT      20-APR-2001; 2000KR-0021070.
FT      12-APR-2001; 2001KR-0019645.
FT      (RYUM/) RYU W.
FT      Lee J, Jeong JK, Cho WY, Yoon GS, Ryu W;
FT      WPI; 2002-019322/03.
FT      Novel prototype hepatitis B virus vector useful in gene therapy for
FT      liver diseases, comprising two novel cis-acting elements essential for
FT      hepatitis B virus genome replication, consisting of alpha and beta
FT      element
FT      Example 2; Page 36-40; 62pp; English.
XX      The invention relates to prototype a hepatitis B virus vector comprising
XX      two novel cis-acting elements essential for hepatitis B virus genome
XX      replication, consisting of an alpha element and a beta element. The
XX      vector is useful for transducing target cells by administering the
XX      into liver tissue of individuals who are chronically infected with
```

```
CC      HBV. The population of target cells are human hepatocytes, and a foreign
CC      strand of host genes as well as viral genes including HBV and hepatitis C
CC      virus, genes encoding tumour suppressors, growth factors, hormones,
CC      cytokines, coagulation factors, and cellular receptors for various
CC      ligands. The vector is also useful in gene therapy for delivering
CC      therapeutic genes to liver cells, and for treating liver diseases and
CC      genetic diseases, chronic HBV infection and for treating metabolic
CC      diseases, such as haemophilia lacking factor VIII or IV expression in
CC      liver. The vector is useful for the expression of heterologous genes in
CC      liver cells. The present sequence is a plasmid, pcMV-HBV/30 which
CC      contains HBV sequences and is used in an experiment demonstrating the
CC      replication competency of wild-type pregenomic RNA expression plasmids.
XX      Sequence 8007 BP; 2022 A; 1980 C; 1824 G; 2181 T; 0 other;
XX
SQ      Query Match          95.5%; Score 1128.2; DB 24; Length 8007;
SQ      Best Local Similarity 97.2%; Pred. No. 0;
SQ      Matches 1148; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
XX
QY      1 ATGGGGCAGAAATCTTCCACCGAGCAATCTCTGGGATTTCTCCGACCAACCAAGTTGAT 60
DB      1031 ATGGGGCAGAAATCTTCCACCGAGCAATCTCTGGGATTTCTCCGACCAACCAAGTTGAT 1090
QY      61 CCAGCCTTCAGAGCAAAACCAACCAATCCAGATTGGAGCTTCATCCCAAGAGACACC 120
DB      1091 CCAGCCTTCAGAGCAAAACCAACCAATCCAGATTGGAGCTTCATCCCAAGAGACACC 1150
QY      121 TGGCCAGAGCCCAAGAGTAGAGCTGAGCATTCGAGCTGGAGGTTCAACCCACCGCAC 180
DB      1151 TGGCCAGAGCCCAAGAGTAGAGCTGAGCATTCGAGCTGGAGGTTCAACCCACCGCAC 1210
QY      181 GGAGGCCCTTTGGGGTGGAGCCCTTGAAGCTCAGGCGATTAACAACAACCTTGGCAGCAAT 240
DB      1211 GGAGGCCCTTTGGGGTGGAGCCCTTGAAGCTCAGGCGATTAACAACAACCTTGGCAGCAAT 1270
QY      241 CGGCTCTGCTTCACCAATGCGCAGTCAAGAAAGGACCTAACCCGCTGCTCAACT 300
DB      1271 CGGCTCTGCTTCACCAATGCGCAGTCAAGAAAGGACCTAACCCGCTGCTCAACT 1330
QY      301 TTGAGAAACACTCATCTCAAGCCATGACAGTAGAATCTCACAACTTTCACCAAACTCTG 360
DB      1331 TTGAGAAACACTCATCTCAAGCCATGACAGTAGAATCTCAACAACTTTCACCAAACTCTG 1390
QY      361 CAAGATCCCAAGTAGAGTGTGTATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB      1391 CAAGATCCCAAGTAGAGTGTGTATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
QY      421 AACCTGTTCGACTACTGTCTCTCCATATGCTCAATCTTCTGAGGATTGGGAGCCCT 480
DB      1451 AACCTGTTCGACTACTGTCTCTCCATATGCTCAATCTTCTGAGGATTGGGAGCCCT 1510
QY      481 GCGCGAAACATGAGAAATCATCATCAGATTCTCAGAACCCCTGCTGCTTACAGGCG 540
DB      1511 GCGCGAAACATGAGAAATCATCATCAGATTCTCAGAACCCCTGCTGCTTACAGGCG 1570
QY      541 GGGTTTTCTTGTGACAAAGATCTCTCAATACCGCAGAGTCTGAGTCTGCTGCTGCTGCTG 600
DB      1571 GGGTTTTCTTGTGACAAAGATCTCTCAATACCGCAGAGTCTGAGTCTGCTGCTGCTGCTG 1630
QY      601 TCTCTCAATTTCTATGGGGGAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB      1631 TCTCTCAATTTCTATGGGGGAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1690
QY      661 TCCATCACTACCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB      1691 TCCATCACTACCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1750
QY      721 CGGCGTTTATCATTTCTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB      1751 CGGCGTTTATCATTTCTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1810
QY      781 CTGACATCAAGGATATGCGCGTGTGCTCTCAATTCAGGATCTTCAACCAACGACG 840
```

```

Db      1811 CTGGACTATCAAGGTATGTTGCCCGTTCCTTAATTCAGGATCCTCAACACGAGC 1870
Qy      841 ACCGGACCATGACGAGCTGCAGACCTCCGCTCAAGAAACCTATGATATCCCTCTGT 900
Db      1871 ACCGGACCATGACGAGCTGCATGACTACTAGCTCAAGAAACCTATGATATCCCTCTGT 1930
Qy      901 TGGTGTACAAACCTTCGATGAGAACTGACACTGTATTCCTCATCCCATCTCTGGGCT 960
Db      1931 TGGTGTACCAACCTTCGATGAGAAATTGACCTGTATTCCTCATCCCATCTCTGGGCT 1990
Qy      961 TTGGGAAAATTCCTATGAGAGTGGGCTCAGCCCGTTTCTCTGGCTCAGTTTACTAGTG 1020
Db      1991 TTGGGAAAATTCCTATGAGAGTGGGCTCAGCCCGTTTCTCTGGCTCAGTTTACTAGTG 2050
Qy      1021 CCATTGCTCAGTGGTCTGTAGAGGCTTCCCGCACTGTTGGCTTTCAGTTATATGAGATG 1080
Db      2051 CCATTGCTCAGTGGTCTGTAGAGGCTTCCCGCACTGTTGGCTTTCAGTTATATGAGATG 2110
Qy      1081 ATGTTGACTGGGGGCCAAGTCTGTACACCATCTTGAAGTCCCTTTTACCGCTGTACCA 1140
Db      2111 ATGTTGATATGGGGGCCAAGTCTGTACACCATCTTGAAGTCCCTTTTACCGCTGTACCA 2170
Qy      1141 ATTTCTTTTGTCTTTGGGCTATACATTAAACCTAATAA 1181
Db      2171 ATTTCTTTTGTCTTTGGGCTATACATTAAACCTAATAA 2211

RESULT 5
AAQ29106
ID      AAQ29106 standard; DNA; 1201 BP.
XX
AC      AAQ29106;
XX
DT      25-MAR-2003 (updated)
DT      24-FEB-1993 (first entry)
XX
XX      13L promoter/S12/core gene.
XX
XX      Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KM      vaccinia virus; Amersia moorei entomopoxvirus; AmbPV; 42 kd; promoter;
KM      NYVAC; recombinant; HBV L; large pre-S antigen; lpsAg; fusion protein;
KM      pre-S region; S12/core; S1; S2; Copenhagen vaccine strain;
XX      vaccinia virus; virulence factor; deletion loci; recipient loci; ss.
XX
OS      Synthetic.
XX
XX      Key
FH      Location/Qualifiers
FT      CDS
FT      32..1201
FT      /*tag= a
FT      /note= "lpsAg gene"
FT      promoter
FT      1..31
FT      /*tag= b
FT      /note= "EBV 42 kd promoter"
FT      misc_RNA
FT      32..355
FT      /*tag= c
FT      /label S1
FT      misc_RNA
FT      356..520
FT      /*tag= d
FT      /label= S2
FT      misc_RNA
FT      521..1198
FT      /*tag= e
FT      /label= S
XX
XX      MO9215672-A1.
XX
XX      17-SEP-1992.
XX
XX      09-MAR-1992; 92WO-US01906.
XX
XX      07-MAR-1991; 91US-0666056.
XX      11-JUN-1991; 91US-0713967.
XX      06-MAR-1992; 92US-0847951.

```

```

XX      (VIRO-) VIROGENETICS CORP.
PA
XX      Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;
PI      Limbach KJ, Norton EK, Paolietti E, Perkins ME, Pincus SB;
PI      Riviére M, Taregila J, Taylor J;
XX
XX      WPI; 1992-331718/40.
DR      P-PSDB; AAR27474.
XX
PT      Vaccine comprises recombinant, attenuated pox-virus - use for
PT      vaccinating against viral infections such as rabies, hepatitis B,
PT      HIV, HSV, EBV, CMV, mumps etc.
XX
PS      Disclosure; Fig 15; 456bp; English.
XX
XX      The sequence given encodes the hepatitis B virus (HBV) L protein
CC      (large pre-S antigen, lpsAg) which is precisely linked to the Amersia
CC      moorei entomopoxvirus (AmbPV) 42 kd promoter. This sequence was used
CC      in the construction of a NYVAC recombinant expressing the HBV gene.
CC      Other HBV genes were also used in the construction. These were HBV M
CC      protein (small pre-S antigen, spsAg) and a fusion protein composed of
CC      the entire pre-S region (S12/core, S1 + S2). Each of these gene
CC      sequences were inserted individually into three different sites of
CC      NYVAC separated by from each other by large regions of vaccinia DNA
CC      containing essential genes. NYVAC is a Copenhagen vaccine strain of
CC      vaccinia virus which has been modified by deletion of six non-essential
CC      regions of the genome encoding known or potential virulence factors.
CC      The deletion loci were engineered as recipient loci for the insertion
CC      of foreign genes. The spacing of the three inserted sequences ensured
CC      that any recombination that did occur would lead to disruption of the
CC      vaccinia genome and would cause unviable vaccinia virus. See also
CC      AAQ35501-864.
CC      (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ      Sequence 1201 BP; 265 A; 350 C; 255 G; 331 T; 0 other;

Query Match      94.6%; Score 1117.2; DB 13; Length 1201;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy      1 ATGGGGGAGAAATCTTTCCACAGCAATCTCTGGATTTCTTCCGACACCGATTGGAT 60
Db      32 ATGGGGGAGAAATCTTTCCACAGCAATCTCTGGATTTCTTCCGACACCGATTGGAT 91
Qy      61 CCAGCCTTGAGAGCAACCAACCAATCCAGATTGGAGCTTCAATCCCAACAGACACC 120
Db      92 CCAGCCTTGAGAGCAACCAACCAATCCAGATTGGAGCTTCAATCCCAACAGACACC 151
Qy      121 TGGCAGACGCCCAACAGTAGAGCTGAGCATTTGAGCTGGGCTTCAACCCGAC 180
Db      152 TGGCAGAGCGCAACAGTAGAGCTGAGCATTTGGGCTTCAACCCGAC 211
Qy      181 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATTAACAACAACCTTCCAGCAAT 240
Db      212 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATTAACAACAACCTTCCAGCAAT 271
Qy      241 CCGCTCTCGCTTCCACCAATTCGCAAGTGAAGGAGCGTCCCGCTGTCTCCACCT 300
Db      272 CCGCTCTCGCTTCCACCAATTCGCAAGTGAAGGAGCGTCCCGCTGTCTCCACCT 331
Qy      301 TTGAGAAACACTATCTCCTCAAGCCATGACATGCAATCCCAACTTTCACCAACTCTG 360
Db      332 TTGAGAAACACTATCTCCTCAGGCGCATGAGTGAATTCACCAACTTTCACCAACTCTG 391
Qy      361 CAAGATCCCAAGTAGAGGTCTGTATTTCTCTGCTGGTGGCTTCCAGTTCAGGAACAGTA 420
Db      392 CAAGATCCCAAGTAGAGGTCTGTATTTCTCTGCTGGTGGCTTCCAGTTCAGGAACAGTA 451
Qy      421 AACCGTTCCGACTCTGTCTCCCATATGTCATATCTTCGAGGATTGGGAGACCT 480
Db      452 AACCGTTCCGACTCTGTCTCCCATATGTCATATCTTCGAGGATTGGGAGACCT 511

```

```

QY 481 GCGCGAGACATGAGAACATCATCATGAGATTCCTAGACCCCTGCTGTTACAGCG 540
DB 512 GCGCTGACATGAGAACATCATCATGAGATTCCTAGACCCCTGCTGTTACAGCG 571
QY 541 GGGTTTCTTGTGACAGAAATCTCTCAATTCGCGAGATCTAGCTGCTGGACT 600
DB 572 GGGTTCTTGTGACAGAAATCTCTCAATTCGCGAGATCTAGCTGCTGGACT 631
QY 601 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGGCAAAATTCGAGTCCCAAC 660
DB 632 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGGCAAAATTCGAGTCCCAAC 691
QY 661 TCCCAATCACTCAACAACCTCTGCTCTGCTCAACTTGTCTGTTATCGCTGATGTC 720
DB 692 TCCCAATCACTCAACAACCTCTGCTCTGCTCAACTTGTCTGTTATCGCTGATGTC 751
QY 721 CGGCTTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 752 CGGCTTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 811
QY 781 CTGGAATCAAGGTATGTTGCGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
DB 812 CTGGAATCAAGGTATGTTGCGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
QY 841 ACGGACATGAGACCTGCTGCACTCTGCTCAAGAACTCTATGATCCCTCTCT 900
DB 872 ACGGACATGAGACCTGCTGCACTCTGCTCAAGAACTCTATGATCCCTCTCT 931
QY 901 TCTGTACAAAACCTTGGATGAGAACTGCACTGTATTCCTCATCCATCATCTGGCT 960
DB 932 TCTGTACAAAACCTTGGATGAGAACTGCACTGTATTCCTCATCCATCATCTGGCT 991
QY 961 TCCGAAAATTCCTATGAGAGTGGGCTGAGCCCGTTCTCTCGGCTCAGTTACTAG 1020
DB 992 TCCGAAAATTCCTATGAGAGTGGGCTGAGCCCGTTCTCTCGGCTCAGTTACTAG 1051
QY 1021 CCATTGTTCAGTGTTCCTAGGGCTTTCCTCCCACTGTGTGCTTCACTTATGATG 1080
DB 1052 CCATTGTTCAGTGTTCCTAGGGCTTTCCTCCCACTGTGTGCTTCACTTATGATG 1111
QY 1081 ATGTGTACTGGGGGCAAGTCTGTAACCAATCTTGAATCCCTTTTACCGCTGTTACA 1140
DB 1112 ATGTGTACTGGGGGCAAGTCTGTAACCAATCTTGAATCCCTTTTACCGCTGTTACA 1171
QY 1141 ATTTCTTTTGTCTTGGTATACATTAA 1170
DB 1172 ATTTCTTTTGTCTTGGTATACATTAA 1201

RESULT 6
AAQ29104
ID AAQ29104 standard; DNA; 1285 BP.
XX
XX AAQ29104;
AC
XX
XX 25-MAR-2003 (updated)
DT 24-FEB-1993 (first entry)
XX
DE lpsAg gene and u promoter.
XX
XX Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KM HBV L; large pre-S antigen; lpsAg; fusion protein; pre-S region;
KM S12/core; S1, S2; Copenhagen vaccine strain; vaccinia virus;
KM virulence factor; deletion loci; recipient loci; ss.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH 116..1285
FT CDS
FT /tag= a
FT /note= "lpsAg gene"
FT promoter 1..115

```

```

FT /tag= b
FT /note= "u promoter"
FT 116..439
FT misc_RNA
FT /tag= c
FT /label S1
FT 440..604
FT /tag= d
FT /label= S2
FT misc_RNA
FT 605..1282
FT /tag= e
FT /label= S

W09215672-A1.
17-SEP-1992.

XX
XX 09-MAR-1992; 92MO-US01906.
XX
XX 07-MAR-1991; 91US-066056.
XX 11-JUN-1991; 91US-0713967.
XX 06-MAR-1992; 92US-0847951.
XX
XX (VIRO-) VIROGENETICS CORP.
XX
XX Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;
PI Limbach KJ, Norton BK, Paolletti B, Perkus ME, Pincus SE;
PI Riviere M, Tartaglia J, Taylor J;
XX
XX WPI; 1992-331718/40.
XX P-PSDB; AAR27472.
XX
XX Vaccine comprises recombinant, attenuated pox-virus - use for
PT vaccinating against viral infections such as rabies, hepatitis B,
PT HIV, HSV, EBV, CMV, mumps etc.
XX
XX Disclosure; Fig 11; 456pp; English.
XX
XX The sequence given encodes the hepatitis B virus (HBV) L protein
CC (large pre-S antigen, lpsAg) which is precisely linked to the combox
CC hemeraghic region (u) promoter. This sequence was used in the
CC construction of a NYVAC recombinant expressing the HBV gene. Other
CC HBV genes were also used in the construction. These were HBV M
CC protein (small pre-S antigen, spsAg) and a fusion protein composed of
CC the entire pre-S region (S12/core, S1 + S2). Each of these gene
CC sequences were inserted individually into three different sites of
CC NYVAC separated by from each other by large regions of vaccinia DNA
CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
CC vaccinia virus which has been modified by deletion of six non-essential
CC regions of the genome encoding known or potential virulence factors.
CC The deletion loci were engineered as recipient loci for the insertion
CC of foreign genes. The spacing of the three inserted sequences ensured
CC that any recombination that did occur would lead to disruption of the
CC vaccinia genome and would cause unviable vaccinia virus. See also
CC AAQ35501-864.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX
XX Sequence 1285 BP; 284 A; 364 C; 272 G; 365 T; 0 other;
SQ

Query Match 94.6%; Score 1117.2; DB 13; Length 1285;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGGGGGCAATCTTTCCACGACATCTCTGGGATTTCTTCCGACACCAAGTTGAT 60
DB 116 ATGGGGGCAATCTTTCCACGACATCTCTGGGATTTCTTCCGACACCAAGTTGAT 175
QY 61 CCAGCTTCAGAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGAGACACC 120
DB 176 CCAGCTTCAGAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGAGACACC 235
QY 121 TGGCCAGAGCCCAACAGGTAGAGCTGAGCATTCGACTGGGGTTACCCCAACCGCAGC 180
DB 236 TGGCCAGAGCCCAACAGGTAGAGCTGAGCATTCGAGCATTCGAGGGCTGGTTTACCCCAACCGCAGC 295

```





```

QY 301 TTGAGAAACATCTCATCTCCAGCATGAGTGAATCCACAATCTTCCACCAACTCTG 360
DB 312 TTGAGAAACATCTCATCTCCAGCATGAGTGAATCCACAATCTTCCACCAACTCTG 371
QY 361 CAAAGATCCCAAGTGAAGAGTCTGTATTTTCCCTGCTGCTGCTCCAGTTTCAAGAACAGTA 420
DB 372 CAAAGATCCCAAGTGAAGAGTCTGTATTTTCCCTGCTGCTGCTCCAGTTTCAAGAACAGTA 431
QY 421 AACCCGTGTCGACATCTGTCTCCCATATGCTCAATCTTCCGAGATGAGGAGCCCT 480
DB 432 AACCCGTGTCGACATCTGTCTCCCATATGCTCAATCTTCCGAGATGAGGAGCCCT 491
QY 481 GCGCGAATCATGAGAACATCATCATGAGATTCCTAGAGACCCCTGCTCTGTTACAGAGG 540
DB 492 GCGCGAATCATGAGAACATCATCATGAGATTCCTAGAGACCCCTGCTCTGTTACAGAGG 551
QY 541 GGGTTTTTCTTTGTGACAAAGATCTCTACAAATCCGACAGATCTAGACTCTGCTGAGT 600
DB 552 GGGTTTTTCTTTGTGACAAAGATCTCTACAAATCCGACAGATCTAGACTCTGCTGAGT 611
QY 601 TCTCAATTTTCTAGGGGGAACTACCGTGTCTTGGCAAAATTCGAGTCCCAACG 660
DB 612 TCTCAATTTTCTAGGGGGAACTACCGTGTCTTGGCAAAATTCGAGTCCCAACG 671
QY 661 TCCAAATCACTCAACCACTCTGCTCTCCACTTGTCTGTTATGCTGATGCTG 720
DB 672 TCCAAATCACTCAACCACTCTGCTCTCCACTTGTCTGTTATGCTGATGCTG 731
QY 721 CGGCGTTTTATCATCTTCTCTTCACTCTGCTGCTATGCTCATCTTCTGTTGTTCT 780
DB 732 CGGCGTTTTATCATCTTCTCTTCACTCTGCTGCTATGCTCATCTTCTGTTGTTCT 791
QY 781 CTGGAATCAAGATATGTTGCCGTTTCTCTTAATCCAGATCTTCAACCAACG 840
DB 792 CTGGAATCAAGATATGTTGCCGTTTCTCTTAATCCAGATCTTCAACCAACG 851
QY 841 ACGGACCATGAGAGCTGACAGACTCTGCTCAAGAACTCTATGATCCCTCTCT 900
DB 852 ACGGACCATGAGAGCTGACAGACTCTGCTCAAGAACTCTATGATCCCTCTCT 911
QY 901 TGTCTGACAAACCTTTCGAGTGAATCTGACCTGTATTCGATCCCATCTGAGCT 960
DB 912 TGTCTGACAAACCTTTCGAGTGAATCTGACCTGTATTCGATCCCATCTGAGCT 971
QY 961 TTCGAAATATCTTATGAGAGTGGGCTTCCGCTTCTCTGCTGCTATCTAGT 1020
DB 972 TTCGAAATATCTTATGAGAGTGGGCTTCCGCTTCTCTGCTGCTATCTAGT 1031
QY 1021 CCATTGTGATGATGCTGTAGAGGCTTTCGCCCACTGTTGCTTCACTTATAGATG 1080
DB 1032 CCATTGTGATGATGCTGTAGAGGCTTTCGCCCACTGTTGCTTCACTTATAGATG 1091
QY 1081 ATGTGTATGAGGGGCAAGTCTG-TACACCATCTTGAATCCCTTTTACCGCTGTAC 1139
DB 1092 ATGTGTATGAGGGGCAAGTCTG-TACACCATCTTGAATCCCTTTTACCGCTGTAC 1151
QY 1140 AATTTCTTTTCTTGGTATATCACTTAAACCTTAATAA 1181
DB 1152 AATTTCTTTTCTTGGTATATCACTTAAACCTTAATAA 1193

```

RESULT 8  
AAH77562  
ID AAH77562 standard; DNA; 3248 BP.

AAH77562;  
19-OCT-2001 (first entry)  
HBV genotype G strain FRI DNA.  
Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBx; HBp1;  
HBsAg; antiHbs; vaccine; genotype G; genotyping; HBsAg; HBsAg; FRI; ds.

```

XX XX Hepatitis B virus.
OS WO200140279-A2.
PN 07-JUN-2001.
PD 20-NOV-2000; 2000WO-BP11526.
PF 03-DEC-1999; 99EP-0870252.
PR 07-DEC-1999; 99US-0169287.
XX (INNO-) INNOGENETICS NV.
XX Stuyver L, Van Geyt C, De Gendt S;
XX WPI; 2001-374785/39.
XX Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX Claim 1; Fig 1; 94pp; English.
XX The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype
CC G. This genotype was found with a high prevalence in patients
CC chronically infected with HBV and residing in Europe and the USA. The
CC invention relates to a fully defined sequence of 3248 nucleotides as
CC given in specification, a sequence with 92% identity to the given
CC sequence, or sequence that is degenerate to the mentioned sequences.
CC These polynucleotides are useful for HBV genotyping. The proteins
CC encoded by the polynucleotides are useful for detecting antibodies in
CC a biological sample. Ligands that bind to the proteins and antibodies
CC directed against the proteins are useful for detecting the proteins
CC and for detecting HBsAg and HBeAg (precore precursor proteins). They
CC are also useful for preparing a vaccine or medicament for treating
CC HBV infections. The present sequence is the complete nucleotide
CC sequence of HBV genotype G strain FRI, which contains four major open
CC reading frames.
XX Sequence 3248 BP; 739 A; 868 C; 700 G; 941 T; 0 other;
SQ

```

Query Match 86.4%; Score 1020.6; DB 22; Length 3248;  
Best Local Similarity 91.6%; Pred. No. 3.1e-310;  
Matches 1080; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```

QY 3 GGGGAGAAATCTTTCACAGCAATCTCTGAGATTTTCCGACCAACAGTGGATCC 62
DB 1103 GGGAAAGAACTTTTCGCGCAGCAATCTCTAGATTTCTTCCGATCACCAGTTGAGCC 1162
QY 63 AGCTTCAGAGCAACACCAACCAATCAGATTGGACTTCAATCCCAACAGAGACCTG 122
DB 1163 AGCATTCAGAGCAAAATACCAACCAATCAGATTGGAGACTTCAATCCCAAAAGAGACCTG 1222
QY 123 GCCAAGCCCAAGAGTGAAGCTGAGCATTCGAGATGGGTTTACCCACCGACCG 182
DB 1223 GCCAAGCCCAAGAGTGAAGCTGAGCATTCGAGATGGGTTTACCCACCGACCG 1282
QY 183 AGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGCAATTAACAACAACCTTGCCAGAAATCC 242
DB 1283 AGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGCAATTCACCTAACAATTTGCGAGCATTC 1342
QY 243 GCTCTGCTTCCACCAATGCGCACTGAGAGGAGCGACTACCCGCTGCTTCCACTTT 302
DB 1343 GCTCTGCTTCCACCAATGCGCACTGAGAGGAGCGACTACCCGCTGCTTCCACTTT 1402
QY 303 GAGAAACATCTACTCTCAAGCCATCAGTGAATCTCCAACTTTTCCACCAACTCTGCA 362
DB 1403 AAGAGACAGTATCTCTCAGGCGCATCAGTGAATCTCTCAGATTTCCACCAAGCTCTACA 1462
QY 363 AGATCCCAAGTGAAGGCTGTATTTCCCTGCTGCTGCTCCACTTCCAGAAAGTAAA 422

```

```

Db      1463  AAATCCCAAGTAGAGGGGCGCTGATTTTCTGCTGGGCTCCAGTTCAAGGATATAGTAA 1522
Qy      423  CCGGTGTCGAGACACTGCTCTCCCATATGTCATTTCTCGAGATTTGGGACCGCTGC 482
Db      1523  CCGGTGTCGAGACACTGCTCTCCCATATGTCATTTCTCGAGATTTGGGACCGCTGC 1582
Qy      483  GCGGAACATGAGAGAACATCATCAGAGATTCTTAGAGACCCCTGCTCGTTTACAGCGGG 542
Db      1583  ACCGAACATGAGAGAACATCATCAGAGATTCTTAGAGACCCCTGCTCGTTTACAGCGGG 1642
Qy      543  GTTTTCTTTGTCAGAAATCTCTCACAAATCCGACAGTCTAGACTCTGCTGGACTTC 602
Db      1643  GTTTTCTTTGTCAGAAATCTCTCACAAATCCGACAGTCTAGACTCTGCTGGACTTC 1702
Qy      603  TCTCAATTTCTGAGGGGAACTACCGGTGCTTGGCAAAATTCGAGTCCCACTC 662
Db      1703  TCTCAATTTCTGAGGGGAACTACCGGTGCTTGGCAAAATTCGAGTCCCACTC 1762
Qy      663  CAATCACTCACCAACTCTGCTCTCCAACTTGTCTGTATTCGTGATGTGTCTGCG 722
Db      1763  CAATCACTCACCAACTCTGCTCTCCAACTTGTCTGTATTCGTGATGTGTCTGCG 1822
Qy      723  GCGTTTATATCATCTTCTCTTTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTTCT 782
Db      1823  GCGTTTATATCATCTTCTCTTTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTTCT 1882
Qy      783  GGAATATCAAGATATGTTGGCCGTTTGCCCTTAATTCAGAGATCTTCAACACACAGAC 842
Db      1883  GGAATATCAAGATATGTTGGCCGTTTGCCCTTAATTCAGAGATCTTCAACACACAGAC 1942
Qy      843  GGAACATGAGAGAGCTGACGACTCTGCTCAAGAGAACCTATATGATCCCTCTGTTG 902
Db      1943  GGAACATGAGAGAGAGCTGACGACTCTGCTCAAGAGAACCTATATGATCCCTCTGTTG 2002
Qy      903  CTGTACAAACCTTTCGATGGAATGCACTGCTATTTCCATCCCATCATCTTGGGCTTT 962
Db      2003  CTGTACAAACCTTTCGATGGAATGCACTGCTATTTCCATCCCATCATCTTGGGCTTT 2062
Qy      963  CGGAATAATTCCTATGAGAGTGGGCTCAGGCCGTTTCCGCTCAGTTTACTAGTGGC 1022
Db      2063  CGGAATAATTCCTATGAGAGTGGGCTCAGGCCGTTTCCGCTCAGTTTACTAGTGGC 2122
Qy      1023  ATTGTTCAGTGTCTCTGAGGCTTTTCCCACTGTTTGGCTTTCAGTTATATGATGAT 1082
Db      2123  ATTGTTCAGTGTCTCTGAGGCTTTTCCCACTGTTTGGCTTTCAGTTATATGATGAT 2182
Qy      1083  GTTGTACTGGGGGCAAGTCTGTACACCATTTTGAATCCCTTTTACCGCTGTTACCAAT 1142
Db      2183  GTTGTATTTGGGGGCAAGTCTGTACACCATTTTGAATCCCTTTTACCGCTGTTACCAAT 2242
Qy      1143  TTTCTTTTGTCTTTGGGATACATTAAACCCCTAATAA 1181
Db      2243  TTTCTTTTGTCTTTGGGATACATTAAACCCCTAATAA 2281

```

RESULT 9  
AAD09091  
ID AAD09091 standard; DNA; 3248 BP.

AC AAD09091;  
DT 04-SEP-2001 (first entry)  
XX Hepatitis B virus FRI strain genotype G DNA.  
DE Hepatitis B virus FRI strain genotype G DNA.  
XX HBV genotype G; precore; core antigen; HBcAg; HBpOL; polymerase; preS1;  
KW envelope protein; preS2; surface antigen; HBsAg; HBX protein; vaccine;  
XX liver disease; hepatitis; liver cancer; circular; cyclic; ds.  
OS Hepatitis B virus.  
XX  
FH Key Location/Qualifiers

```

FT      CDS 1..675
FT      /tag= a
FT      /product= "PreCore/HBcAg core protein"
FT      /transl_except= (pos:4..6, aa:Xaa)
FT      /note= "Xaa corresponds to in-frame stop codon"
FT      /tag= b
FT      /note= "PreCore protein DNA"
FT      /tag= c
FT      /note= "HBcAg core protein DNA"
FT      /tag= d
FT      /note= "Core insert peptide DNA"
FT      /tag= e
FT      /product= "HBpOL protein"
FT      /tag= f
FT      /note= "Corresponds to the end position (3248) of the
FT      circular HBV DNA"
FT      /tag= g
FT      /note= "Corresponds to the start position (1) of the
FT      circular HBV DNA"
FT      CDS 1071..2270
FT      /tag= h
FT      /product= "PreS1-PreS2-HBsAg envelope proteins"
FT      /tag= i
FT      /product= "PreS1 envelope protein"
FT      /note= "Does not include stop codon"
FT      /tag= j
FT      /product= "PreS2 envelope protein"
FT      /note= "Does not include stop codon"
FT      /tag= k
FT      /product= "Hepatitis B surface antigen (HBsAg) protein"
FT      /tag= l
FT      /product= "HBX protein"
FT      /transl_except= (pos:3244..4, aa:Ala-Cys)
FT      /note= "The nucleotides 2809-3246 and 2-25 encode
FT      amino acids 1-146 and 147-153, respectively of the
FT      HBX protein (AAE04712)"
FT      /tag= m
FT      /product= "HBX protein fragment (AAE04714)"
FT      /note= "Does not include start and stop codons"
FT      /tag= n
FT      /product= "HBX protein fragment (AAE04713)"
FT      /note= "Does not include start codon"
FT      /partial
FT      CDS 2..25
FT      /tag= n
FT      /product= "HBX protein fragment (AAE04713)"
FT      /note= "Does not include start codon"
FT      /partial
XX      W0200138498-A2.
XX      31-MAY-2001.
XX      21-NOV-2000; 2000MO-US32108.
XX      24-NOV-1999; 99US-0167206.
XX      (PHAR-) PHARMASSET INC.
XX      (INNO-) INNOGENETICS NV.
XX      Stuyver L, Schnazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M,
XX      Rosbau R;
XX      WPI; 2001-367676/38.

```

DR P-ESDB; AAE04707, AAE04708, AAE04709, AAE04710, AAE04711, AAE04712,  
DR AAE04713, AAE04714.

PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,  
PT polypeptides encoded by nucleic acids, useful for preparing vaccine to  
PT treat or prevent the hepatitis B virus genotype G infection in a  
PT subject -

Claim 3; Fig 6; 84pp; English.

CC The present sequence is hepatitis B virus (HBV) strain FRL,  
CC genotype G DNA encoding PreCore/Core protein, HBsPol, envelope  
CC (PreS1, PreS2 and surface antigen HBsAg) and HBx proteins.  
CC HBV genotype G nucleic acids and polypeptides are useful for diagnosing  
CC and prognosing and treating infections caused by HBV genotype G. They  
CC can be used in a vaccine to treat or prevent HBV genotype G infection.  
CC The HBV genotype G derived nucleic acids and antibodies are useful for  
CC detecting HBV genotype G in a sample or diagnosis of HBV genotype G  
CC infection. The presence of HBV genotype G statistically correlates with  
CC the presence of liver damage and/or liver cancer in the subject. The  
CC HBV genotype G core insert peptide encoding nucleic acid is useful for  
CC designing monitoring assays to study and predict the evolution of  
CC anti-HBe and anti-HBc antibodies and HBeAg (genotype G e antigen) in  
CC patients infected with HBV. The antibodies or antigens of HBV genotype G  
CC are useful for identifying a stage of liver disease caused by HBV  
CC genotype G.

**SQ** Sequence 3248 BP; 739 A; 868 C; 700 G; 941 T; 0 other;

Query Match	86.4%	Score 1020.6	DB 22	Length 3248
Best Local Similarity	91.6%	Pred. No. 3.1e-310		
Matches 1080, Conservative	0	Mismatches 99	Indels 0	Gaps 0

QY	3	GGGGCAGAAATCTTTTCCACCAACAATCCTCTGGAGATTCCTTTCCCGACCAACAGTTGGATTC	62
Db	1103	GGGAAAGAACTTTTTCGGCAGCAAAATCCTTAAGATTCCTTCCGATCAACAGTTGGAACTC	1162
QY	63	AGCCTTCAGAGCAAAACCAACAATTCAGATTGGGACTTCAATCCCAACAAGACACTG	122
Db	1163	AGCAATTCAGAGCAAAATTCACAATCCGATTGGGACTTCAATCCCAAAAAGAACCTTG	1222
QY	123	GCCAGACGCCAACAAGGTAGAGCTGAGACTTGGACTGGAGTTGGGGTTTCAACCCACGCAAG	182
Db	1223	GCCAGAGGCCAACAAGGTAGAGGTGGAGCTTAAGAACCGGGTTTCAACCTTCACACGG	1282
QY	183	AGGCTTTTGGGGGTGGAGGCTTCAGGCTCAGGGCATAACACAACCTTGCAGCAAAATCC	242
Db	1283	AGGCTTTTGGGGGTGGAGGCTTCAGGCTCAGGGCACAATAACAATTGTGCAGAGATTC	1342
QY	243	GCCCTCGCTTCCACAATGGCCAGTACGAGAAAGCAGCTAACCCCGCTGTCCACCTT	302
Db	1343	GCCCTCGCTTCCACAACCAATGCTCAGTACGGAGGAGCAGCTAACCTTCACCACT	1402
QY	303	GAGAAACACTCATCTCAAGCCATGAGTGAATCTCCAATTTTCCACAAACTTGTCA	362
Db	1403	AAGAGACAGTATCTCTCAGGCGCAGTGAATCTTAAGACATTTCCACCAAGCTTACA	1462
QY	363	AGATCCCAAGTAGAGAGTCTGTATTTCTCTGTGTGGCTTCAGTTCAAGAACGTAA	422
Db	1463	AAATCCCAAGTAGAGGCGCTGTATTTTCTGTGTGTGGCTTCAGTTCAAGGAATGTAA	1522
QY	423	CCCGTTCCGACTGTCTCTCCACTATGTCACATCTTTCAGAGGATTTGGGGACCCGTC	482
Db	1523	CCCGTTCCGACTATGCTCTCTCAATTCGTCAATCTTCTCAGAGATTTGGGGACCCGTC	1582
QY	483	GCGGAACATGAGAACATCATCAGGATTCCTAGAACCCCTGCTCGTTTACAGGCGGG	542
Db	1583	ACCGAACATGAGAACATCATCAGGATTCCTAGAACCCCTGCTCGTTTACAGGCGGG	1642
QY	543	GTTTTTCTGTGTGACAGAACTCTCAATACCCGACAGTCTAAGCTGTGTGTGACTTC	602
Db	1643	GTTTTTCTGTGTGACAGAACTCTCAATACCCGACAGTCTAAGCTGTGTGTGACTTC	1702

Oy		603	TCTCAATTTTCTAGGGGGAAC	CTACC	GGTGTCTTGCGCAAAATTCGAGTCCCAACTC	662
Dd		1703	TCTCAATTTTCTAGGGGGAAGTGC	CGTGTCTTGCGCAAAATTCGAGTCCCAACTC	1762	
Oy		663	CATTCACCTCACCAACTCTCTGT	CTCTCCAATTGTCCTGGTATATGCGTGAATGTGTCTGCG	722	
Dd		1763	CATTCACCTCACCAACTCTCTGT	CTCTCCAATTGTCCTGGTATATGCGTGAATGTGTCTGCG	1822	
Oy		723	GCGTTTTATCATCTTCTCTTCAT	CTCGTGTCTAATGCTCATCTTCTTGTGTGTTCTCT	782	
Dd		1823	GCGTTTTATCATATTCCTTCCTT	CATCTCGTGTCTAATGCTCATCTTCTTGTGTGTTCTCT	1882	
Oy		783	GGACTATATCAAGATATGTTCGCC	CTTTGTCTCTTAATTCGAGATCTTACAACACACACAC	842	
Dd		1883	GGACTATATCAAGATATGTTCGCC	CTTTGTCTCTTAATTCGAGATCTTACAACACACACAC	1942	
Oy		843	GGGACCAATGCGAGCCTGCACGA	CTCGTGTCAAAGAACCTTAATGATCCCTCTGTG	902	
Dd		1943	GGGACCTGCAAAACCTGCACGAC	TCTCTCAAGGCACTTAATGATCCCTCAATGTG	2002	
Oy		903	CTGTACAAAACCTTCGATGGAAC	CTGCACCTGTATTCGATCCCATCATCTGTGGCTTT	962	
Dd		2003	CTGTACAAAACCTTCGATGGAAC	CTGCACCTGTATTCGATCCCATCATCTGTGGCTTT	2062	
Oy		963	CGAAAAATTCCTTAGAGAGTGG	CGCTCAGCCCCGTTTCTCTCGTCACTTACTAGTCCC	1022	
Dd		2063	CGAAAAATTCCTTAGAGAGTGG	CGCTCAGCCCCGTTTCTCTCGTCACTTACTAGTCCC	2122	
Oy		1023	ATTGTTCAGGTGTGATAGGGCTT	TGCCACATCTTGTGCTTCACTTAATATGATGAT	1082	
Dd		2123	ATTGTTCAGGTGTGATAGGGCTT	TGCCACATCTTGTGCTTCACTTAATATGATGAT	2182	
Oy		1083	GTTGTACTGGGGGCAAGTCTGT	ACACCATCTTGAGTCCCTTTTACCGCTTTACCAAT	1142	
Dd		2183	GTTGTACTGGGGGCAAGTCTGT	ACACCATCTTGAGTCCCTTTTACCGCTTTACCAAT	2242	
Oy		1143	TTTCTTTTGTCTTGGGTATACAT	TTAACCTTAATAA 1181		
Dd		2243	TTTCTTTTGTCTTGGGTATACAT	TTAACCTTAATAA 2281		
<b>RESULT 10</b>						
ID	AAH77564	standard; DNA; 1203 BP.				
XX	AAH77564;					
XX	AC					
DT	19-OCT-2001	(first entry)				
XX	HBV pres/S gene.					
DE	HBV pres/S gene.					
KM	Hepatitis B virus; HBV, preCore, Core; pres1, pres2, HBS; HBx; HBPol;					
KM	HBsAg; antiviral; vaccine; genotype G; genotyping; HbsAg; HBeAg;					
XX	pres/S; ds.					
OS	Hepatitis B virus.					
PN	WO200140279-A2.					
PD	07-JUN-2001.					
PF	20-NOV-2000; 2000MO-EPI1526.					
PR	03-DEC-1999; 99EP-0870252.					
PR	07-DEC-1999; 99US-0169287.					
PA	(INNO-) INNOGENETICS NV.					
PI	Stuyver L, Van Geys C, De Gendt S;					
DR	WPI; 2001-374785/39.					
PT	Novel isolated and/or purified hepatitis B virus polypeptide and					

PT polynucleotide sequences that are phylogenetically different from HBV  
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and  
PT therapy -

PS Claim 4; Fig 3; 94pp; English.

CC The invention relates to the complete nucleic acid sequence of a new  
CC human hepatitis B virus (HBV) genotype, provisionally named genotype  
CC G. This genotype was found with a high prevalence in patients  
CC chronically infected with HBV and residing in Europe and the USA. The  
CC invention relates to a fully defined sequence of 3248 nucleotides as  
CC given in specification, a sequence with 92% identity to the given  
CC sequence, or sequence that is degenerate to the mentioned sequences.  
CC These polynucleotides are useful for HBV genotyping. The proteins  
CC encoded by the polynucleotides are useful for detecting antibodies in  
CC a biological sample. Ligands that bind to the proteins and antibodies  
CC directed against the proteins are useful for detecting the proteins  
CC and for detecting HBcAg and HBeAg (precursor proteins). They  
CC are also useful for preparing a vaccine or medicament for treating  
CC HBV infections. The present sequence is the complete coding  
CC sequence of the HBV pres/S gene

**SQ** Sequence 1203 BP; 256 A; 353 C; 262 G; 328 T; 4 other;

Query Match	85.5%	Score 1010.2	DB 22	Length 1203
Best Local Similarity	91.5%	Pred. No. 3.4e-307		
Matches 1069	Conservative	0	Mismatches 99	Indels 0
				Gaps 0

Oy	3	GGGGGAGAAATCTTTTCCACGAAATCTCTGGGATTTCTTCCCGACACAGTTGGATTC	62
Db	36	GGGAAGAAGACTTTTCCACGAAATCTCTAGATTTCTTCCCGATCAGAGTTGGACC	95
Oy	63	AGCCTTCAGAGCAACACCAACATCCAGATTGGGACTTCAATCCCAACAAAGACACTG	122
Db	96	AGCATTCAGAGCAAAATACCAACAAATCCAGATTGGGACTTCAATCCCAAAAAGACCCTTG	155
Oy	123	GCCAGAGCCCAACAGATAGAGAGCTGGAGCATTTGGACTGGGGGTTACACCCACCGCAGG	182
Db	156	GCCAGNAGCAACAGATAGAGAGTTGGAGCTTATGGACCAGGAGTTCACCCCTCCACAGG	215
Oy	183	AGGCTTTTGGGGGTGGAGCCCTCAGGCTCAGGGGCATTAACACAAACCTTGCCAGCAATCC	242
Db	216	AGGCTTTTGGGGGTGGAGCCCTCAGCTTCAGGGGCACACTAACAACTTTGGCAGCAGATCC	275
Oy	243	GCTTCCTGCTTCCACCATTGCGCACTCAGAAAGGCAAGCTTACCCGCTGTCTCCACTTT	302
Db	276	GCTTCCTGCTTCCACCATTGCTCACTCAGGGAGGAGGCTTATCCCATCTCTCCACACT	335
Oy	303	GAGAAACACTCATCTCTCAAGCCATCAGTGGAACTCCCAACTTTCCACCAACTCTGCA	362
Db	336	AAGAGACAGTATCTCTCAGGCCATCAGATGGAACTCTACAGATTCCACCAAGCTCTACA	395
Oy	363	AGATCCCAAGAGTGAAGGTCTGTATTTTCCCTGCTGTGGCTTCAGTTCAGAAACGTAAA	422
Db	396	AAATCCCAAGTCAAGGGCTGTATTTTCTCTGCTGTGGCTTCAGTTCAGGAGTGTGAA	455
Oy	423	CCCTGTCCGACTACGTCTCTCCCATATCCGCATTTCTTCGAGAGATTGGGAGCCCTGC	482
Db	456	CCCTGTCCGACTATTTGCTCTCAACTCTCGTCAATCTTCTCAGAGATTGGGAGCCCTGCG	515
Oy	483	GCGGAACATGAGAACATCACATCAGAGATTCCTAGAGCCCTGCTCGTGTTCACAGCGGG	542
Db	516	ACCGAACAATGAGAACATCACATCAGAGATTCCTAGAGCCCTGCTCGTGTTCACAGCGGG	575
Oy	543	GTTTTTTCTTGTGACAAAGATCTCTCAATACCGCAGACTCTAGACTGTGTGGACTTC	602
Db	576	GTTTTTTCTTGTGACAAAGATCTCTCAATACCGCAGACTCTAGACTGTGTGGACTTC	635
Oy	603	TCTCAATTTTCTAAGGGGGAATACCGTGTGCTTGGCCAAAATTCGCAAGTCCCAACCTC	662
Db	636	TCTCAATTTTCTAAGGGGGAAGTCCCGTGTGTCTTGGCCTAATTCGCAAGTCCCAACCTC	695
Oy	663	CAATCACTCACCACCTCTGTCTTCCAACTTGTCTGTGATTCGCTGGATGTCTGCG	722

Db	696	CAATCACTACCAATCTCTGTCTCTCAACTGTTCCTGGCTATCGCTGATGTGTCTGGC	755
Qy	723	GCGTTTATCATCTTCTCTTTCATCTCGTGTCTATAGCTTCATCTTGTGTGTTCTTCT	782
Db	756	GCGTTTATCATATATCTCTCTTTCATCCGCTGTCTATGCTCATCTTCTGTGTGTTCTTCT	815
Qy	783	GGACTATCAAGATATGTTCGCCGTTTGTCTCTTAATTCAGAAATCTTCAACCAACAGCAC	842
Db	816	GGACTATCAAGATATGTTCGCCGTTTGTCTCTGTATTCAGAAATCTTCAACCAACATAC	875
Qy	843	GGAGCAATGCAGAGCCTGCAGACTCCTGTCTCAAGAACTATATGATCCCTCTGTG	902
Db	876	GGAGCACTGCAGAAACCTGCAGACTCCTGTCTCAAGGAATCTATATATCCCTCATGTG	935
Qy	903	CTGTACAAAACCTTTCGAGTGAATCTGACCTGTATTTCCATCCCATCATCTGTGGCTTT	962
Db	936	CTGTACAAAACCTTTCGAGCGGAAATTCACCTGTATTTCCATCCCATCGACTGTGGGCTTT	995
Qy	963	CGGAAAATTTCTATAGGAGTGTGGGCTCAAGCCGTTTCTCTGGCTCAAGTTACTAGTGC	1022
Db	996	CGCAAAATATCTTATAGGAGTGTGGGCTCAAGTCCGTTTCTCTGGCTCAAGTTACTAGTGC	1055
Qy	1023	ATTATGTCAGTGTTCGTGATGGGCTTCCCACTGTTGTGCTTCAAGTTATATGATGAT	1082
Db	1056	ATTATGTCAGTGTTCGTGATGGGCTTCCCACTGTCGTGGCTTCAAGTTATATGATGAT	1115
Qy	1083	GTTGTACTGGGGGCAAGTCTGTACACATCTTGATGCTCTTTTACCGCTTTACCAAT	1142
Db	1116	GTTGTACTGGGGGCAAGTCTGTACACATCTTGATGCTCTTTTATACCGCTTTACCAAT	1175
Qy	1143	TTTCTTTTGTCTTTGGGTATCATTTAA	1170
Db	1176	TTTCTTTTGTCTTTGGGTATCATTTAA	1203

RESULT 11

ID AAQ05378 standard; DNA; 3835 BP.

AC AAQ05378;

DT 10-DEC-1990 (first entry)

Sequence of hepatitis B genome.

tumorigenesis; vaccine; hepatitis; ds; kw

OS Hepatitis B virus.

PN JP02163089-A

PD 22-JUN-1990.

PF 15-DEC-1988; 88JP-0315097.

PR 15-DEC-1988; 88JP-0315097.

PA (DAUC ) DAIICHI PHARM CO LTD.

DR WPI; 1990-235346/31

PT HBV-DNA sequence and transgenic animal - used for screening for

XX

XX

CC transgenic animal expressing HBV DNA, useful in studying the mechanism

Sequence 3835 BP; 825 A; 1048 C; 870 G; 1092 T; 0 other;

```

Query Match      85.1%; Score 1005; DB 11; Length 3835;
Best Local Similarity 90.7%; Pred. No. 2.8e-305;
Matches 1071; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 1 ATGGGGGAGAAATCTTCCACCAAGCATCTCTGGGAGTTCTTCCGACCAAGTTGGAT 60
DB 1481 ATGGGAGCAAAATCTTCTGTTCCCAATCCTCTGGGATCTTCTCCATACCAAGTTGGAC 1540
QY 61 CCAGCCTTCAGAGCAAAACCAACAATTCAGATTGGGACTTCAATCCCAACAAGACACC 120
DB 1541 CCGGGTTGGAGCCCAACTCACCAATCCGATTGGGACTTCAACCCCAACAAGATCAT 1600
QY 121 TGGCCAGACGCCAACAAGTAGAGCTGAGCATTCGGACTGGGGTTCAACCCCAAGCAC 180
DB 1601 TGGCCAGAGCAAAATCAGGTAGAGGGAGCATTCGGGCTCAGGGTTCAACCCCAAGCAC 1660
QY 181 GGAGGGCTTTGGGGGGGAGCCCTCAGGCTCAGGGGATTAACAACAACCTTGGCAGCAAT 240
DB 1661 GGCGGCTTTTGGGGGGGAGCCCGCAGGCTCAGGGGGTTGAACAACCGTGCAGATAGCA 1720
QY 241 CCGCCTCTGCTTCCACCAATCGCCAGAGAAAGGACCTTACCCGCTGTCTCCACCT 300
DB 1721 CCTCCTCCTGCTCCACCAATCGGAGTACAGAGAAAGACCTTACTCTCTCCACCT 1780
QY 301 TTGAGAAACATCTCATCTCAAGCATGAGTGAACCTCCACAACCTTCCACAACCTCTG 360
DB 1781 CTAAAGAGACATCTCATCTCAGGCGATGAGTGAACCTCCACAACCTTCCACAAGCTCTG 1840
QY 361 CAAGATCCCAAGATGAGAGTGTGTATTTCCCTGCTGTGGCTCCAGTTCCAGAAACATGA 420
DB 1841 CTAGAGCCCAAGATGAGGGGCTTATCTTCTGCTGTGGCTCCAGTTCCAGAAACATGA 1900
QY 421 AACCCCTGCTCCAGTACTGTCTCTCCATATGCTCATCTTTCAGAGATGGGGAGCCCT 480
DB 1901 AACCTGCTCCAGTACTGTCTCTCCATATGCTCATCTTTCAGAGATGGGGAGCCCT 1960
QY 481 GCGGGAACATGAGAAACATCAATCAGATTCCTAGAGACCCCTGCTCTGTTCAGAGCG 540
DB 1961 GCACCGAATGAGAAACATCAATCAGATTCCTAGAGACCCCTGCTCTGTTCAGAGCG 2020
QY 541 GGGTTTTTCTTGTGACAAAGATCTCTCAACAATCCGAGAGTCTAGACTCTGTGTGACT 600
DB 2021 GGGTTTTTCTTGTGACAAAGATCTCTCAACAATCCGAGTCTAGACTCTGTGTGACT 2080
QY 601 TCTCTCAATTTCTAGGGGGAACTACCGTGTCTTGGCCAAAATTCGAGTCCCAAC 660
DB 2081 TCTCTCAATTTCTAGGGGGAGACCCACGTCCTGGCCAAAATTCGAGTCCCAAC 2140
QY 661 TCCAACTCACTCAACACTCTGTCTCTCAACTGTCTGTATTCGCTGATGTGTCTG 720
DB 2141 TCCAACTCACTCAACACTCTGTCTCTCAACTGTCTGTATTCGCTGATGTGTCTG 2200
QY 721 CGGGCTTTTATCTCTCTCTTCACTCTGCTGTATTCGCTCATCTCTTGTGTGTTCTT 780
DB 2201 CGGGCTTTTATCTCTCTCTTCACTCTGCTGTATTCGCTCATCTCTTGTGTGTTCTT 2260
QY 781 CTGACATTCAGAGTATGTGCGCGTTGTCTCTAATTCGAGATCTTCAACCAACGAC 840
DB 2261 CTGACATTCAGAGTATGTGCGCGTTGTCTCTAATTCGAGATCTTCAACCAACGAC 2320
QY 841 ACGGAGCATGAGAGCCCTGACGACTCTGCTCAAGAAACCTCATATGATCCCTCTCTGT 900
DB 2321 ACGGAGCATGAGAGCCCTGACGAGTCTGCTCAAGAAACCTCATATGATCCCTCTCTGT 2380
QY 901 TGTCTGACAAAACCTTCGAGTGAACCTGACCTGTATTCGATCCCATCTCTGAGGT 960
DB 2381 TGTCTGACAAAACCTTCGAGTGAACCTGACCTGTATTCGATCCCATCTCTGAGGT 2440
QY 961 TTGGAGAAATCTCTATNGAGAGTGGGCTTCAGCCGCTTCTCTGCTAGTTACTAGTG 1020
DB 2441 TTGGAGAAATCTCTATNGAGAGTGGGCTTCAGCCGCTTCTCTGCTAGTTACTAGTG 2500
QY 1021 CCATTGTTCAGTGTGCTGAGGCTTTCGCCCACTGTTGGCTTTCAGTTATATGATG 1080

```

```

DB 2501 CCATTGTTCAGTGTGCTGAGGCTTTCGCCCACTGTTGGCTTTCAGTTATATGATG 2560
QY 1081 ATGTGTACTGGGGGCCAAGTCTGTACACCAATCTGAGTCCCTTTTACCGGTATACA 1140
DB 2561 ATGTGTATTTGGGGGCCAAGTCTGTACACCAATCTGAGTCCCTTTTACTCTATATACA 2620
QY 1141 ATTTCTTTTGTCTTGGGTATACATTTAAACCTATATAA 1181
DB 2621 ATTTCTTTTGTCTTGGGTATACATTTAAACCTATATAA 2661

RESULT 12
AAD14316
ID AAD14316 standard; DNA; 4084 BP.
XX
AC AAD14316;
XX
DT 06-NOV-2001 (first entry)
XX
DE Hepatitis B virus (HBV) 1.28 genome.
XX
KW Hepatitis B virus; HBV; altered sensitivity; agent; detection;
KW PCR primer; Hepatitis B surface antigen; Hbsag; ds.
XX
OS Hepatitis B virus.
PN W0200157244-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-AU00098.
XX
PR 03-FEB-2000; 2000US-0179948.
XX
PA (MELB-) MELBOURNE HEALTH.
PA (PENN-) PENN STATE RES FOUND.
PI Delaney W, Locarnini SA, Chen RYM, Bartholomeusz A, Isom H;
XX
DR WPI; 2001-496926/54.
XX
PT Detecting hepatitis B virus variant with altered sensitivity to agent,
PT comprises infecting genetic construct containing replication competent
PT genome to cells, contacting cells with agent and detecting replication
PT of variant -
XX
PS Example 7; Fig 5A; 110bp; English.
XX
CC The invention relates to a method of detecting variant hepatitis B virus
CC (HBV) which exhibits altered sensitivity to agents. The method involves
CC infecting a genetic construct containing a replication competent amount
CC of the genome from variant HBV contained in or fused to a baculovirus
CC genome; contacting cells with the agent to be tested; culturing cells
CC under conditions sufficient for the variant HBV to replicate, express
CC genetic sequences, and/or assemble, and/or release viral particles; and
CC determining replication of variant HBV using viral-component-detection
CC means. The method is useful for detecting variant HBV which exhibits
CC altered sensitivity to agents. The present sequence is the HBV 1.28
CC genome.
XX
SQ Sequence 4084 BP; 920 A; 1105 C; 914 G; 1145 T; 0 other;

Query Match      84.8%; Score 1001.8; DB 22; Length 4084;
Best Local Similarity 90.5%; Pred. No. 2.9e-304;
Matches 1069; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 ATGGGAGCAATCTTTCACAGCAATCTCTGGATTTCTTCCGACCAAGTTGAT 60
DB 1436 ATGGGAGCAATCTTTCCTGTTCCCAACCTCTGGGATTTCTTCCGATCAGTTGAC 1495
QY 61 CCAGCTTCAGAGAAACCAACAATCAGATTGGGACTTCATCCCAACAAGGACAC 120

```

```
Db 1496 CTGCAATTCGAGCCCAACTCAATCAATCGAATTGGAGCTTCAACCCCATCAAGACAC 1555
Qy 121 TGGCCAGACCCCAACAGGTAGAGCTGGAGCAATTCGGACTGGGGTTACCCACCGCAC 180
Db 1556 TGGCCAGACCCCAACAGGTAGAGCTGGAGCAATTCGGAGCTGGGGTTACCCACCGCAC 1615
Qy 181 GGAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGCACTTAAACAAACCTTGGCCAGCAAT 240
Db 1616 GGGGATTTTGGGGTGAAGCCCTCAGGCTCAGGGCACTTAAACAAACAGTGTCAACAAAT 1675
Qy 241 CCGGCTCTCTCTTCCACCAATCGCCAGTCAGAAAGCAGCCCTACCCGCTGTCCACCT 300
Db 1676 CTTCTCTCTCTCTTCCACCAATCGCCAGTCAGAAAGCAGCCCTACCCATCTCTCCACCT 1735
Qy 301 TTGAGAAACATCTCTCTCAAGCCATGAGGAACTCCCAACTTTCACCAAACTCTG 360
Db 1736 CTAAAGAACATCTCTCTCAAGCCATGAGGAACTCCCAACTTTCACCAAACTCTG 1795
Qy 361 CAAGATCCCAAGAGTGAAGGTCTGTATTTCCCTGCTGGTGGCTCAAGTTCAGAAACAGTA 420
Db 1796 CAGATCTCCAGAGTGAAGGTCTGTATTTCCCTGCTGGTGGCTCAAGTTCAGAAACAGTA 1855
Qy 421 AACCTGTCTCGACTCTCTCTCTCCATATCTGCAATCTTCTCAGAGATTTGGGACCT 480
Db 1856 AACCTGTCTCGAATATTTGCTCTCACATCTCTGCAATCTCAGAGAGTGGGGACCT 1915
Qy 481 GCGGGAAACATGGAACATCAATCAGATCTCTAGAGCCCTGCTCGTTTACAGCG 540
Db 1916 GTGACGAACATGGAACATCAATCAGATCTCTAGAGCCCTGCTCGTTTACAGCG 1975
Qy 541 GGGTTTTCTTGTGACAAAGATCTCAATACCGAGAGTCTAGACTCGTGTGACT 600
Db 1976 GGGTTTTCTTGTGACAAAGATCTCAATACCGAGAGTCTAGACTCGTGTGACT 2035
Qy 601 TCTCTCAATTTTCTAGGGGAACTACCGGTGTCTTGGCCAAATTTGCGAGTCCCAACC 660
Db 2036 TCTCTCAATTTTCTAGGGGAACTACCGGTGTCTTGGCCAAATTTGCGAGTCCCAACC 2095
Qy 661 TCCAAATCACTACCAACTCTCTCTCTCCAACTTGTCTGTATTCCTGATGTGCTG 720
Db 2096 TCCAAATCACTACCAACTCTCTCTCTCCAACTTGTGTATTCCTGATGTGCTG 2155
Qy 721 CGGGCTTTTATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 2156 CGGGCTTTTATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2215
Qy 781 CTGAGTATCAAGGTATGTGCGCGTTGTCTCTAATTCAGAGTCTTCAACCAACAGC 840
Db 2216 CTGAGTATCAAGGTATGTGCGCGTTGTCTCTAATTCAGAGTCAACCAACAGC 2275
Qy 841 ACCGGACATGCAAGCTGTGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 2276 ACCGGACATGCAAACTGTGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2335
Qy 901 TGCTGTACAAAACTTGGATGAAATGCACTGTATTCCTCATCCATCCATCTCTGGGCT 960
Db 2336 TGCTGTACAAAACTTGGATGAAATGCACTGTATTCCTCATCCATCTCTGGGCT 2395
Qy 961 TTGGGAAATTCATGAGGAGTGGGCTCAGCCCGTTCTCTGCTGCTCAAGTTTACTAGT 1020
Db 2396 TTCCCAAAATTCATGAGGAGTGGGCTCAGCTCGTTCTCTTGGCTCAGTTTACTAGT 2455
Qy 1021 CCATTGTGCTAGTGTCTGTAGGGCTTCCCACTGTATTTGGCTTTCAGTTATATAGATG 1080
Db 2456 CCATTGTGCTAGTGTCTGTAGGGCTTCCCACTGTATTTGGCTTTCAGTTATATAGATG 2515
Qy 1081 ATGTGTACTGGGGCCAAAGTCTGTACACATCTTGAATCCCTTTTACCGCTGTACCA 1140
Db 2516 ATGTGTACTGGGGCCAAAGTCTGTACACATCTGAGATCCCTTTTATCCGCTGTACCA 2575
Qy 1141 ATTTTCTTTTGTCTTGGGTATATCAATTTAAACCTTAATAA 1181
Db 2576 ATTTTCTTTTGTCTTGGGTATATCAATTTAAACCTTAATAA 2616
```

```
RESULT 13
AADD14317
ID AADD14317 standard; DNA; 4496 BP.
XX
AC AADD14317;
XX
DT 06-NOV-2001 (first entry)
XX
DE Hepatitis B virus (HBV) 1.5 genome.
XX
KM Hepatitis B virus; HBV; altered sensitivity; agent; detection;
XX PCR primer; Hepatitis B surface antigen; HBsAg; ds.
OS Hepatitis B virus.
XX
PN WO200157244-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001MO-AU00098.
XX
PR 03-FEB-2000; 2000US-0179948.
XX
PA (MELB-) MELBOURNE HEALTH.
XX (PENN-) PENN STATE RES FOUND.
XX
PI Delaney W, Locarnini SA, Chen RYM, Bartholomew A, Isom H;
XX
DR WPI; 2001-496926/54.
XX
PT Detecting hepatitis B virus variant with altered sensitivity to agent,
PT comprises infecting genetic construct containing replication competent
PT genome to cells; contacting cells with agent and detecting replication
PT of variant.
XX
PS Example 7; Fig 5B; 110bp; English.
XX
CC The invention relates to a method of detecting variant hepatitis B virus
CC (HBV) which exhibits altered sensitivity to agents. The method involves
CC infecting a genetic construct containing a replication competent amount
CC of the genome from variant HBV contained in or tested to a baculovirus
CC genome; contacting cells with the agent to be tested; culturing cells
CC under conditions sufficient for the variant HBV to replicate, express
CC genetic sequences, and/or assemble, and/or release viral particles; and
CC determining replication of variant HBV using viral component-detection
CC means. The method is useful for detecting variant HBV which exhibits
CC altered sensitivity to agents. The present sequence is the HBV 1.5
CC genome.
XX
SQ Sequence 4496 BP; 999 A; 1229 C; 1011 G; 1257 T; 0 other;
XX
Query Match 84.8%; Score 1001.8; DB 22; Length 4496;
Best Local Similarity 90.5%; Pred. No. 3.1e-304;
Matches 1069; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
Qy 1 ATGGGCGAATCTTTCCACCAAGATCTCTGGGATTTCTTCCGACCAAGGATGGAT 60
Db 1848 ATGGGCGAATCTTTCTGTTCCCAACCTCTGGGATTTCTTCCCAATCAAGTGGAG 1907
Qy 61 CCAGCTTTCAGAGCAACCAACCAATCCAGATTGGGACTTTCATCCCAACAGACACC 120
Db 1908 CTGCAATTCGAGCCCAACTCAACCAATCCAGATTGGGACTTTCATCCCAACAGACACC 1967
Qy 121 TGGCCAGACCCCAACAGGTAGAGCTGGAGCTTGGAGCTGGGGTTACCCACCGCAC 180
Db 1968 TGGCCAGACCCCAACAGGTAGAGCTGGAGCAATTCGGGCTCAAGGCTCACCCCTCCACAC 2027
Qy 181 GGAGGCTTTTGGGGTGAAGCCCTCAGGCTCAGGGCACTTAAACAAACCTTGGCCAGCAAT 240
Db 2028 GGGGATTTTGGGGTGAAGCCCTCAGGCTCAGGGCACTTATTAACCAAGTCAACAAAT 2087
```

```

QY 241 CCGCCTCTGCTTCCACCAATGCGAGTGAAGAGGACGTAACCCGCTGTCTCCACT 300
DB 2088 CCTCTCTCCGCTCCCAACCAATGCGAGTGAAGAGGACGTAACCTCTCTCCACT 2147
QY 301 TTGAGAAACATCATCTCTCAAGCCATGAGTGAAGTCTCAACAATTTCCACCAACTCTG 360
DB 2148 CTAAGAGACATCATCTCTCAAGCCATGAGTGAAGTCTCAACAATTTCCACCAAGTCTG 2207
QY 361 CAAGATCCAGAGTGAAGTCTGTATTTCCCTGCTGAGTCTCCAGTTCAAGAAACAGTA 420
DB 2208 CAGATCCAGAGTGAAGTCTGTATTTCCCTGCTGAGTCTCCAGTTCAAGAAACAGTA 2267
QY 421 AACCTGTTCCGACTACTGTCTCTCCATATGCTCAATCTTCTCGAAGATTGGGACCT 480
DB 2268 AACCTGTTCCGAAATATGCTCTCTCAATCTGCTCAATCTTCCGCGAGAGTGGGACCT 2327
QY 481 GGGCGGACATGAGAAACATCATATGAGATTTCTTAGAACCCCTGCTGCTTACAGGG 540
DB 2328 GTAGCGAACATGAGAAACATCATATGAGATTTCTTAGAACCCCTGCTGCTTACAGGG 2387
QY 541 GGGTTTTTCTGTTGACAAGAAATCTCACAATACCGCAGAGTCTAGACTCGTGTGAGCT 600
DB 2388 GGGTTTTTCTGTTGACAAGAAATCTCACAATACCGCAGAGTCTAGACTCGTGTGAGCT 2447
QY 601 TCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 660
DB 2448 TCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 2507
QY 661 TCCATCATCTCAACCACTCTCTCTCTCAATCTGCTGAGTCTGCTGAGTGTCTG 720
DB 2508 TCCATCATCTCAACCACTCTCTCTCTCAATCTGCTGAGTCTGCTGAGTGTCTG 2567
QY 721 CGGCGTTTATCATCTCTCTCTCTCAATCTGCTGAGTCTGCTGAGTGTCTG 780
DB 2568 CGGCGTTTATCATCTCTCTCTCTCAATCTGCTGAGTCTGCTGAGTGTCTG 2627
QY 781 CTGACTATCAAGATATGTTGCCGTTTCTCTCAATCTGAGATCTTCAACCAACAGC 840
DB 2628 CTGACTATCAAGATATGTTGCCGTTTCTCTCAATCTGAGATCTTCAACCAACAGC 2687
QY 841 ACGGACCATGAGAGCTCTGACAGTCTCTCTCAAGAACTCTATGATCTCTCTG 900
DB 2688 ACGGACCATGAGAAACCTGACAGTCTCTCTCAAGAGCACTATGATCTCTCTG 2747
QY 901 TSCGTACAAAACCTTCGATGAGAACTGACCTGATCCCATCCATCATCTCTGAGT 960
DB 2748 TSCGTACAAAACCTTCGATGAGAACTGACCTGATCCCATCCATCATCTCTGAGT 2807
QY 961 TTCGAAATATCTATGAGAGTGGGCTGAGCCGTTTCTCTGCTGAGTCTTACTAGT 1020
DB 2808 TTCGAAATATCTATGAGAGTGGGCTGAGCCGTTTCTCTGCTGAGTCTTACTAGT 2867
QY 1021 CCATTTGTTCAAGTGTCTGATGAGGCTTTTCCCACTGTTTGGCTTCAATATGAGAT 1080
DB 2868 CCATTTGTTCAAGTGTCTGATGAGGCTTTTCCCACTGTTTGGCTTCAATATGAGAT 2927
QY 1081 ATGTGTACTGGGGGCAAGTCTGTAACACATCTGAGCCCTTTTACCGCTTACCA 1140
DB 2928 ATGTGTATGGGGGCAAGTCTGTAACACATCTGAGCCCTTTTACCGCTTACCA 2987
QY 1141 ATTTCTTTTCTTTGGGTATATTAATTAACCTTAATAA 1181
DB 2988 ATTTCTTTTCTTTGGGTATATTAATTAACCTTAATAA 3028

```

## RESULT 14

```

AAT36611
ID AAT36611 standard; cDNA; 1200 BP.
XX
AC AAT36611;
XX
XX 25-MAR-2003 (updated)
DT 22-OCT-1996 (first entry)

```

```

XX DE Hepatitis B virus pre S1/pre S2/S gene (adr-1).
XX KW Hepatitis C virus; HCV; hepatitis B virus; HBV; fusion protein;
XX KW immunisation; vaccine; infection; core protein; adr-1; ds.
XX OS Hepatitis B virus.
XX FH
XX FT misc_difference 151..153
XX FT /tag= a
XX FT /transl_except= 151..153,aa:His
XX FT misc_difference 163..165
XX FT /tag= b
XX FT /transl_except= 161..163,aa:Ala
XX FT misc_difference 273..275
XX FT /tag= c
XX FT /transl_except= 273..275,aa:Val
XX PN
XX PD WO9610997-A1.
XX PD 18-APR-1996.
XX PF 05-OCT-1995; 95WO-US13552.
XX PR 06-JUN-1995; 95US-0467859.
XX PR 05-OCT-1994; 94US-0318248.
XX PA (ABOL-) ABOLION INC.
XX PA (GENO-) GEN HOSPITAL CORP.
XX PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
XX PI Coney LR, Pachuk CJ, Tokushige K, Wakita T, Wands J;
XX PI Zurawski VR;
XX DR WPI; 1996-209642/21.
XX DR P-PSDB; AAR93801.
XX PT Nucleic acid encoding hepatitis B and C virus fusion proteins - or
XX PT incomplete hepatitis C virus genome, are useful in vaccines for
XX PT prevention or treatment of HBV and HCV infections
XX PS
XX PS Example 1; Page 32-34; 53pb; English.
XX CC A new fusion protein comprises a hepatitis B virus (HBV) S gene
XX CC protein (AAT36611) coupled to amino acids 1-69, 1-70 or 1-154 of the
XX CC hepatitis C virus (HCV) core protein (AAT28348).
XX CC The nucleic acid has the coding sequence linked to a CMV promoter,
XX CC RSV enhancer, polyadenylation sequence and opt. the 5'-UTR of HCV.
XX CC (updated on 25-MAR-2003 to correct PA field.)
XX SQ
XX SQ Sequence 1200 BP; 255 A; 351 C; 265 G; 329 T; 0 other;
XX
XX Query Match 84.7%; Score 1000.6; DB 17; Length 1200;
XX Best Local Similarity 91.1%; Pred. No. 3.6e-304;
XX Matches 1063; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 1 ATGGGACGAATCTTCCACGACGATCTCTGAGATTTCTTCCGACGACGATTGGAT 60
DB 34 ATGGGAGCAATCTTCTGTTCTCCCAATCTCTGGGATTTCTTCCGATCACCAGTTGGAC 93
QY 61 CCAGCCTTCAGAGCAACACCAACCAATCCAGATTGGGACTTCATCCCAAGGACACC 120
DB 94 CCGGCTTCAGAGCAACCAACCAATCCAGATTGGGACTTCATCCCAAGGATCA 153
QY 121 TGGCCAGACGCCCAACAGATGAGCTGAGACTTGGAGTTCAGCCGACCGAC 180
DB 154 TGGCCAGAGAAATCAAGATGAGGCGGAGACTTGGGCGCAGGTTCACTCCACACAC 213
QY 181 GGAGGCTTTTGGGGTGAGCCCTCAGGCTCAGGGCATATACAAACCTTGGCAGCAAT 240
DB 214 GCGGCTTTTGGGGTGAGCCCTCAGGCTCAGGGCATATATGACAAACAGTGCAGACAGC 273

```

```
QY 241 CCGCCTCTCTCTCCACCAATCCGCACTGAGAAAGGAGCCTACCCGCTGTCTCCACT 300
DB 274 CTTCTCTCTCTCTCCACCAATCCGCACTGAGAAAGGAGCCTACCTCTCTCTCCACT 333
QY 301 TTGAGAAACATCTCTCTCAAGCCTGAGGAACTCCCAACTTTTCCCAACTCTG 360
DB 334 CTAGAGACAGTCATCTCTCAAGCCATGAGGAACTCCCAACTTTTCCCAAGCTCTG 393
QY 361 CAAGATCCCAAGAGAGAGCTGTATTTCCCTGCTGGTGGCTCAGTTAGAGACAGTA 420
DB 394 CTAGATCCCAAGAGAGAGGAGGCTATATTTCTCTGCTGGTGGCTCAGTTAGAGACAGTA 453
QY 421 AACCTGTCTCGAATCTGTCTCTCCCAATCTGTCAATCTTTCTGAGAGATTGGGACCT 480
DB 454 AACCTGTCTCGAATCTGTCTCTCCCAATCTGTCAATCTTTCTGAGAGACTGGGACCT 513
QY 481 GCGCGGAACATGAGAAATCATCATCGAGATTCTTAGAGCCCTGCTGCTTTAGAGCG 540
DB 514 GCACCGAACAATGAGAGACAAATCAGAGATTCTTAGAGCCCTGCTGCTTTAGAGCG 573
QY 541 GGGTTTTCTTGTGACAAAGATCTCAATACCGAGAGTGAAGCTGGTGGACT 600
DB 574 GGGTTTTCTTGTGACAAAGATCTCAATACCGAGAGTGAAGCTGGTGGACT 633
QY 601 TCTCTCAATTTTCTAGGGGAACTACCGGTGTCTTGGCCAAATTTGCAAGTCCCAACC 660
DB 634 TCTCTCAATTTTCTAGGGGAGACACCGGTGTCTTGGCCAAATTTGCAAGTCCCAACC 693
QY 661 TCCAAATACCTCAACCTCTCTCTCTCAACTTGTCTGGTATCGCTGGATGTCTG 720
DB 694 TCCAAATACCTCAACCTCTCTCTCTCAACTTGTCTGGTATCGCTGGATGTCTG 753
QY 721 CGCGTTTTTATCATCTCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 754 CGCGTTTTTATCATCTCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
QY 781 CTGAGACTATCAAGGATATGTCGCCGTTTGTCTCTTAAATTCAGAGATTTCAACCCAGC 840
DB 814 CTGAGACTATCAAGGATATGTCGCCGTTTGTCTCTTAAATTCAGAGATTTCAACCCAGC 873
QY 841 ACGGAGACATGACAGGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB 874 ACGGAGACATGACAGGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 933
QY 901 TGTGTATCAAAACCTTGGATGAGAACTGCACTGTATTCCTATCCCATCTCTGGACT 960
DB 934 TGTGTATCAAAACCTTGGATGAGAACTGCACTGTATTCCTATCCCATCTCTGGACT 993
QY 961 TTGGGAAATTCCTATGAGAGTGGGCTCAAGCCGTTTCTCTGGCTCAGTTTACTAGTG 1020
DB 994 TTGGGAAATTCCTATGAGAGTGGGCTCAAGCCGTTTCTCTGGCTCAGTTTACTAGTG 1053
QY 1021 CCATTTTCTGAGTGTCTGATGAGGCTTTCCCGCACTGTTGGCTTTAGTTATATGAGTG 1080
DB 1054 CCATTTTCTGAGTGTCTGATGAGGCTTTCCCGCACTGTTGGCTTTAGTTATATGAGTG 1113
QY 1081 ATGTTAGTATGAGGAGCAAGCTGTATACCACTTTGAGTCCCTTTTACCGCTTTACA 1140
DB 1114 ATGTTAGTATGAGGAGCAAGCTGTATACCACTTTGAGTCCCTTTTACCGCTTTACA 1173
QY 1141 ATTTTCTTTTGTCTTTGGGTATACATT 1167
DB 1174 ATTTTCTTTTGTCTTTGGGTATACATT 1200
```

## RESULT 15

```
AA23282
ID AA23282 standard; DNA; 9325 BP.
XX
XX AA23282;
XX
XX
DT 31-JAN-2000 (first entry)
XX
```

```
DE DNA sequence of plasmid pTHBV-d.
XX
XX Hepatitis B virus; HBV; recombinant; pol gene; X gene; surface antigen;
KM liver; anti-viral; anti-tumor; gene therapy; single-gene defect;
KM genetic disorder; familial hypercholesterolemia; neoplastic gene;
KM ornithine transcarbamylase deficiency; ss.
XX
OS Synthetic.
OS Hepatitis b virus.
XX
XX US5981274-A.
XX
XX 09-NOV-1999.
XX
XX 18-SEP-1996; 96US-0715808.
XX
XX 18-SEP-1996; 96US-0715808.
XX
XX (CHAI/) CHAISOMCHIT S.
XX (CHAN/) CHANG L.
XX (TYRR/) TYRRELL D L J.
XX
XX Chang L, Chaiemchit S, Tyrrell DLJ;
PI WPI; 1999-63330/54.
XX
XX Recombinant hepatitis B virus genome containing heterologous gene
XX PT sequences useful for treating liver infections -
XX
XX Example 1; Columns 39-48; 53bp; English.
XX
XX The invention relates to a recombinant hepatitis B virus genome (HBV)
XX CC that comprises heterologous gene sequences which express at least one
XX CC functional heterologous gene product. A host cell transfected with a
XX CC recombinant HBV genome comprising pol gene sequences, X gene sequences
XX CC and surface antigen gene (preS1/preS2/S gene) sequences and heterologous
XX CC gene sequences can be used to express at least one functional
XX CC heterologous gene product. The invention also provides a method for
XX CC encapsidating a recombinant HBV genome. The recombinant HBV genomes are
XX CC useful for the expression of functional heterologous gene products in
XX CC liver cells. The vectors can be used for anti-viral, anti-tumor and/or
XX CC gene therapy and particularly for the correction of inherited single-gene
XX CC defects. Human genetic disorders which can be treated by expression of
XX CC missing or mutant genes in the liver are familial hypercholesterolemia
XX CC and ornithine transcarbamylase deficiency. Primary tumors of the liver
XX CC may benefit from the expression of anti-neoplastic genes in the liver.
XX CC Existing retroviral vectors and other animal viruses which are used to
XX CC deliver foreign genes are not liver-specific with regard to their
XX CC infection or expression unlike hepatitis B viral vectors. Human hepatitis
XX CC B virus can be delivered through the circulation so there is no
XX CC requirement for tissue culture for ex vivo liver-directed gene therapy.
XX CC The present sequence represents the DNA sequence of the plasmid pTHBV-d
XX CC which comprises HBV sequences.
XX
XX SQ Sequence 9325 BP; 2227 A; 2448 C; 2132 G; 2518 T; 0 other;
XX
XX Query Match 84.7%; Score 1000.2; DB 20; Length 9325;
XX Best Local Similarity 90.4%; Pred. No. 1.5e-303;
XX Matches 1068; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2003, 01:24:40 ; Search time 72 Seconds  
(without alignments)  
7239.909 Million cell updates/sec

Title: US-09-821-877-1  
Perfect score: 1181  
Sequence: 1 atggggcagaatcttccac.....tacattaaccctaataa 1181

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1117.2	94.6	1201	1	US-08-105-483-218 Sequence 218, App
2	1117.2	94.6	1201	1	US-08-709-209-218 Sequence 218, App
3	1117.2	94.6	1201	1	US-08-458-101-218 Sequence 218, App
4	1117.2	94.6	1285	1	US-08-105-483-215 Sequence 215, App
5	1117.2	94.6	1285	1	US-08-709-209-215 Sequence 215, App
6	1117.2	94.6	1285	1	US-08-458-101-215 Sequence 215, App
7	1088.2	92.1	2348	3	US-08-480-173A-42 Sequence 42, Appl
8	1088.2	92.1	2348	3	US-08-484-408A-42 Sequence 42, Appl
9	1000.6	84.7	1200	3	US-08-854-531-5 Sequence 5, Appl
10	1000.6	84.7	1200	3	PCT-US95-13552-5 Sequence 5, Appl
11	1000.2	84.7	9325	2	US-08-715-808-2 Sequence 2, Appl
12	1000.2	84.7	9659	2	US-08-715-808-6 Sequence 6, Appl
13	988.2	83.7	3220	6	5196194-15 Patent No. 5196194
14	985.8	83.5	7463	2	US-08-715-808-13 Sequence 13, Appl
15	811.2	68.7	3182	4	US-08-890-735C-1 Sequence 11, Appl
16	810.8	68.7	846	4	US-09-247-890-11 Sequence 11, Appl
17	810.8	68.7	846	4	US-09-724-969-11 Sequence 11, Appl
18	810.8	68.7	846	4	US-09-724-852-11 Patent No. 5164485
19	767.2	65.0	846	6	5164485-1 Sequence 5, Appl
20	767.2	65.0	6371	2	US-08-715-808-5 Sequence 12, Appl
21	767.2	65.0	6371	2	US-08-715-808-12 Sequence 12, Appl
22	764.8	64.8	3221	2	US-08-715-808-14 Sequence 14, Appl
23	764.8	64.8	3221	2	US-08-715-808-1 Sequence 14, Appl
24	756.4	64.0	846	4	US-09-247-890-9 Sequence 9, Appl
25	756.4	64.0	846	4	US-09-724-969-9 Sequence 9, Appl
26	756.4	64.0	846	4	US-09-724-852-9 Sequence 9, Appl
27	756.2	64.0	3220	6	5196194-11 Patent No. 5196194

28	752.6	63.7	822	3	US-08-075-520A-14 Sequence 14, Appl
29	744.6	63.0	3215	2	US-09-719-528A-1 Sequence 1, Appl
30	731.8	62.0	893	4	US-08-500-914A-1 Sequence 1, Appl
31	693	58.7	5618	3	US-08-799-556-1 Sequence 1, Appl
32	693	58.7	5618	4	US-09-570-546-1 Sequence 1, Appl
33	672.4	56.9	801	4	US-09-311-784A-15 Sequence 15, Appl
34	652.2	55.2	681	5	PCT-US96-10602-13 Sequence 13, Appl
35	650.8	55.1	678	3	US-08-075-520A-26 Sequence 26, Appl
36	647.8	54.9	845	1	US-08-378-011A-2 Sequence 2, Appl
37	643.4	54.5	817	1	US-08-378-011A-4 Sequence 4, Appl
38	637	53.9	3504	2	US-08-760-797A-2 Sequence 2, Appl
39	637	53.9	3504	2	US-08-760-797A-4 Sequence 4, Appl
40	637	53.9	3504	3	US-08-932-929B-2 Sequence 2, Appl
41	637	53.9	3504	4	US-08-932-929B-4 Sequence 4, Appl
42	633	53.6	1194	5	PCT-US96-10602-5 Sequence 5, Appl
43	625	52.9	681	2	US-08-500-914A-7 Sequence 7, Appl
44	625	52.9	681	2	US-08-500-914A-9 Sequence 9, Appl
45	623.4	52.8	681	4	US-09-471-573A-1 Sequence 1, Appl

## ALIGNMENTS

```
RESULT 1
US-08-105-483-218
; Sequence 218, Application US/08105483
; Patent No. 5494807
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,483
; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 218:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-105-483-218

Query Match          94.6%; Score 1117.2; DB 1; Length 1201;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

1 ATGGGGCAGAACTCTTCCACGACGACATCTCTGGGATTCTTCCGACGACGATTGAT 60
|||||
```

Db 32 ATGGGGGAGAAATCTTTCCACGCAATCTCTGGGATTTCTTCCGACCAACAGTTGGAT 91  
Qy 61 CCAAGCTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAACAGAGACCC 120  
Db 92 CCAAGCTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAACAGAGACCC 151  
Qy 121 TGGCCAGAGCCCAACAGAGTAGAGAGTAGAGCATTTGGAGCTTGGGGTTTACCCCAACCCGAC 180  
Db 152 TGGCCAGAGCCCAACAGAGTAGAGAGTAGAGCATTTGGGGCTGGGGTTTACCCCAACCCGAC 211  
Qy 181 GAGAGCTTTTGGGGGTGAGAGCCCTCAGAGCTTACAGGCAATACAAACCTTGGCCAGCAAT 240  
Db 212 GAGAGCTTTTGGGGGTGAGAGCCCTCAGAGCTTACAGGCAATACAAACCTTGGCCAGCAAT 271  
Qy 241 CCGCTCTCTGCTTCCACCAATTCGAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
Db 272 CCGCTCTCTGCTTCCACCAATTCGAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331  
Qy 301 TTGAGAAACATCATCTCTCAAGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Db 332 TTGAGAAACATCATCTCTCAAGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391  
Qy 361 CAAGATCCCAAGAGTAG 420  
Db 392 CAAGATCCCAAGAGTAG 451  
Qy 421 AACCTGTTCCGAGTACATGCTCTCCATATCGTCAATCTTCTGAGAGATTTGGGGAGAGCT 480  
Db 452 AACCTGTTCCGAGTACATGCTCTCCATATCGTCAATCTTCTGAGAGATTTGGGGAGAGCT 511  
Qy 481 GCGGGAACATGAGAAATCATCATCAGAGATTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
Db 512 GCGGGAACATGAGAAATCATCATCAGAGATTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 571  
Qy 541 GGGTTTTCTTTTGAAGAAATCTCTCAATATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
Db 572 GGGTTTTCTTTTGAAGAAATCTCTCAATATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 631  
Qy 601 TCTCTCAATTTTCTAGAGAGAGAACTACCGTGTCTTGGCCAAATTTGGAGTCTTGGAGTCT 660  
Db 632 TCTCTCAATTTTCTAGAGAGAGAACTACCGTGTCTTGGCCAAATTTGGAGTCTTGGAGTCT 691  
Qy 661 TCCATCACTACCAACCTCTCTGTCTTCAATCTTGTCTGTGTTATCTGTGAGTGTCTGT 720  
Db 692 TCCATCACTACCAACCTCTCTGTCTTCAATCTTGTCTGTGTTATCTGTGAGTGTCTGT 751  
Qy 721 CGGGTTTTATCATCTTCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 752 CGGGTTTTATCATCTTCTCTTCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811  
Qy 781 CTGAGATATCAAGGATATGTTGCGCGTTTGTCTCTAATTTCCAGAGATCTTCAACAGCAGC 840  
Db 812 CTGAGATATCAAGGATATGTTGCGCGTTTGTCTCTAATTTCCAGAGATCTTCAACAGCAGC 871  
Qy 841 ACGGAGCATGAGAGAGCTGACAGCTCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Db 872 ACGGAGCATGAGAGAGCTGACAGCTCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 931  
Qy 901 TGTGTAACAAACCTTGGAGTAGAACTGCACTGTATTCCTATCCCATCATCTCTGGAGCT 960  
Db 932 TGTGTAACAAACCTTGGAGTAGAACTGCACTGTATTCCTATCCCATCATCTCTGGAGCT 991  
Qy 961 TTGGGAAATTTCTATGAGAGTAG 1020  
Db 992 TTGGGAAATTTCTATGAGAGTAG 1051  
Qy 1021 CCAATTTGTCAGTGTCTGTAAGGCTTTCGCCCACTGTTTGGCTTTTCAAGTTATATGAGATG 1080  
Db 1052 CCAATTTGTCAGTGTCTGTAAGGCTTTCGCCCACTGTTTGGCTTTTCAAGTTATATGAGATG 1111  
Qy 1081 ATGTTGTAAGTGGGGCCCAAGTCTGTACACATCTTGGAGTCCCTTTTAAACGCTGTATACCA 1140  
Db 1112 ATGTTGTAAGTGGGGCCCAAGTCTGTACACATCTTGGAGTCCCTTTTAAACGCTGTATACCA 1171

Qy 1141 ATTTCTTTTGTCTTTGGTATACATTAA 1170  
Db 1172 ATTTCTTTTGTCTTTGGTATACATTAA 1201

RESULT 2  
US-08-709-209-218  
Sequence 218, Application US/08709209  
Patent No. 5762938

GENERAL INFORMATION:  
APPLICANT: Paoli, Enzo  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
NUMBER OF INVENTION: STRAIN  
NUMBER OF SEQUENCES: 462  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,209  
FILING DATE: 21-AUG-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/105,483  
FILING DATE: 12-AUG-1993  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 218:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1201 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-709-209-218

Query Match 94.6%; Score 1117.2; DB 1; Length 1201;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ATGGGGGAGAAATCTTTCCACGCAATCTCTGGGATTTCTTCCGACCAACAGTTGGAT 60  
Db 32 ATGGGGGAGAAATCTTTCCACGCAATCTCTGGGATTTCTTCCGACCAACAGTTGGAT 91  
Qy 61 CCAAGCTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAACAGAGACCC 120  
Db 92 CCAAGCTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAACAGAGACCC 151  
Qy 121 TGGCCAGAGCCCAACAGAGTAGAGAGTAGAGCATTTGGAGCTTGGGGTTTACCCCAACCCGAC 180  
Db 152 TGGCCAGAGCCCAACAGAGTAGAGAGTAGAGCATTTGGGGCTGGGGTTTACCCCAACCCGAC 211  
Qy 181 GAGAGCTTTTGGGGGTGAGAGCCCTCAGAGCTTACAGGCAATACAAACCTTGGCCAGCAAT 240  
Db 212 GAGAGCTTTTGGGGGTGAGAGCCCTCAGAGCTTACAGGCAATACAAACCTTGGCCAGCAAT 271  
Qy 241 CCGCTCTCTGCTTCCACCAATTCGAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

272 CCGCTCCGCTCCCAACCAATGCGCAGAGAGGAGGACCCGCTGCTCCACCT 331  
301 TTGAGAAACATCATCTCTAAGCATGAGTGAATCTCCACAACTTCCACCAACTG 360  
332 TTGAGAAACATCATCTCTAAGCATGAGTGAATCTCCACAACTTCCACCAACTG 391  
361 CAAGATCCAGAGTGAAGGTCTGTATTTCCCTGCTGAGTCCAGTTCAAGAACAGTA 420  
392 CAAGATCCAGAGTGAAGGTCTGTATTTCCCTGCTGAGTCCAGTTCAAGAACAGTA 451  
421 AACCTGTTCCGACTCTGTCTCTCCCATATGTCATCTTTCAGAGATTGGGACCT 480  
452 AACCTGTTCCGACTCTGTCTCTCCCATATGTCATCTTTCAGAGATTGGGACCT 511  
481 GCGCGAATGAGAAACATCATCAGAGATCCAGAGTCCAGAGTCCAGAGTCCAGAG 540  
512 GCGCGAATGAGAAACATCATCAGAGATCCAGAGTCCAGAGTCCAGAGTCCAGAG 571  
541 GCGGTTTCTGTGTGAAGAAATCTCAGAAATCCAGAGTCCAGAGTCCAGAGTCCAGAG 600  
572 GCGGTTTCTGTGTGAAGAAATCTCAGAAATCCAGAGTCCAGAGTCCAGAGTCCAGAG 631  
601 TCTCTCAATTTCTAGAGGAGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660  
632 TCTCTCAATTTCTAGAGGAGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 691  
661 TCCATGACTCAACCACTCTGCTCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
692 TCCATGACTCAACCACTCTGCTCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751  
721 CGGCGTTTATCATCTTCTCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
752 CGGCGTTTATCATCTTCTCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811  
781 CTGACTATCAAGATGATGTTGCCGTTTCTCTCTATTCAGAGATCTTCAACCAACGAC 840  
812 CTGACTATCAAGATGATGTTGCCGTTTCTCTCTATTCAGAGATCTTCAACCAACGAC 871  
841 ACGGAGCAATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
872 ACGGAGCAATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931  
901 TCTCTGACAAACCTTTCGAGATGAGAACTGACCTGTATTCCTCATCTCTGCTGCTGCT 960  
932 TCTCTGACAAACCTTTCGAGATGAGAACTGACCTGTATTCCTCATCTCTGCTGCTGCT 991  
961 TTGCGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTGCTGCTGCTGCTGCTG 1020  
992 TTGCGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTGCTGCTGCTGCTGCTG 1051  
1021 CCATTTGTTCAAGTGTGCTGAGGCTTCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
1052 CCATTTGTTCAAGTGTGCTGAGGCTTCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1111  
1081 ATGTGTACTGAGGAGGCAAGTCTGTACACCATCTTGAAGTCCCTTTTACCGCTGTTACA 1140  
1112 ATGTGTACTGAGGAGGCAAGTCTGTACACCATCTTGAAGTCCCTTTTACCGCTGTTACA 1171  
1141 ATTTCTTTTGTCTTTGGGTATACATTTAA 1170  
1172 ATTTCTTTTGTCTTTGGGTATACATTTAA 1201

## RESULT 3

US-08-458-101-218  
Sequence 218, Application US/08458101  
Patent No. 5765599

## GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo  
APPLICANT: Perkins, Marion E.  
APPLICANT: Taylor, Jill  
APPLICANT: Tartaglia, James

APPLICANT: No. 5765599/en, Elizabeth K.  
APPLICANT: Riviere, Michel  
APPLICANT: de Taisne, Charles  
APPLICANT: Limbach, Keith J.  
APPLICANT: Johnson, Gerard P.  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
APPLICANT: Audomert, Jean-Christophe Francis  
APPLICANT: Gettig, Russell Robert  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
TITLE OF INVENTION: STRAIN  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,101  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2740  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 218:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1201 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-458-101-218  
Query Match 94.6%; Score 1117.2; DB 1; Length 1201;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
1 ATGGGAGCAATCTTCCACGACGATCTCTGGGATCTTCCGACCAAGTTGAT 60  
32 ATGGGAGCAATCTTCCACGACGATCTCTGGGATCTTCCGACCAAGTTGAT 91  
61 CCAGCTTTCAGAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGGACAC 120  
92 CCAGCTTTCAGAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGGACAC 151  
121 TGGCCAGAGCCCAACAGAGTGAAGTGGAGCTTGGAGTCTGGAGTCAACCCGAC 180  
152 TGGCCAGAGCCCAACAGAGTGAAGTGGAGCTTGGAGTCTGGAGTCAACCCGAC 211  
181 GAGGCTTTTGGGAGGAGGCTCAGGCTCAGGCAATACCAAACTTGGCAGCAAT 240  
212 GAGGCTTTTGGGAGGAGGCTCAGGCTCAGGCAATACCAAACTTGGCAGCAAT 271  
241 CGGCTCTGCTTCCCAATTCAGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
272 CGGCTCTGCTTCCCAATTCAGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 331  
301 TTGAGAAACATCATCTCTAAGCATGAGTGAATCTCCACAACTTCCACCAACTG 360  
332 TTGAGAAACATCATCTCTAAGCATGAGTGAATCTCCACAACTTCCACCAACTG 391  
361 CAAGATCCAGAGTGAAGGTCTGTATTTCCCTGCTGAGTCCAGTTCAAGAACAGTA 420

```
|||||
Db 392 CAAGATCCAGAGTGAAGAGCCTGTATTTCCCTGCTGTGGCTCCAGTTCAGAGACAGTA 451
421 AACCCGTTCCGACTCTGTCTCTCCATATCCGATCTTTCCGAGATTTGGGACCC 480
452 AACCCGTTCCGACTCTGTCTCTCCATATCCGATCTTTCCGAGATTTGGGACCC 511
481 GCGCGAACAATGAGAAACATCAGATCAGATTCCTAGAGACCCCTGCTGATTACAGCG 540
512 GCGCTGAACATGAGAAACATCAGATCAGATTCCTAGAGACCCCTGCTGATTACAGCG 571
541 GGGTTTTCTTTGTTGACAAATCTTCAATAACCGAGTCTAGACTGCTGGTGACT 600
572 GGGTTCTTTCTTTGTTGACAAATCTTCAATAACCGAGTCTAGACTGCTGGTGACT 631
601 TCTCTCAATTTTCTAGGGGGAACACACCGTGTCTTGGCCAAAATTTGGAGTCCCAAC 660
632 TCTCTCAATTTTCTAGGGGGAACACACCGTGTCTTGGCCAAAATTTGGAGTCCCAAC 691
661 TCCAACTACTCAACCACTCTCTGCTCAACTGTCTGATTATCGCTGAGTGTCTG 720
692 TCCAACTACTCAACCACTCTCTGCTCAACTGTCTGATTATCGCTGAGTGTCTG 751
721 CGGCGTTTATCATCTTCTCTTCATCTGCTGCTATGCTCATCTTCTTTGTTCTT 780
752 CGGCGTTTATCATCTTCTCTTCATCTGCTGCTATGCTCATCTTCTTTGTTCTT 811
781 CTGACATCAAGATATGTTGCCGTTTGTCTCTATATTCAGAGATTTCAACACAGC 840
812 CTGACATCAAGATATGTTGCCGTTTGTCTCTATATTCAGAGATTTCAACACAGC 871
841 ACAGGACATGAGAGCTGCAAGACTCTGCTCAAGAACTGTATGTATCCCTCTGT 900
872 ACAGGACATGAGAGCTGCAAGACTCTGCTCAAGAACTGTATGTATCCCTCTGT 931
901 TGTGTACAAACCTTGGATGAACTGACCTGTATTCCTATCCATCCATCATCTGGGCT 960
932 TGTGTACAAACCTTGGATGAACTGACCTGTATTCCTATCCATCCATCATCTGGGCT 991
961 TTGGGAAATTTCCATAGGAGTGGGCTCAGCCCGTTTCTCCGCTCAGTTTACTATG 1020
992 TTGGGAAATTTCCATAGGAGTGGGCTCAGCCCGTTTCTCCGCTCAGTTTACTATG 1051
1021 CCATTTGTTAGTGTCTGTAGAGCTTTCCGCCACTGTTGGCTTTCAGTTATATGATG 1080
1052 CCATTTGTTAGTGTCTGTAGAGCTTTCCGCCACTGTTGGCTTTCAGTTATATGATG 1111
1081 ATGTTGACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTACCGCTGTACCA 1140
1112 ATGTTGATTTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTACCGCTGTACCA 1171
1141 ATTTTCTTTGTTGTTGGGTATATCATTTAA 1170
1172 ATTTTCTTTGTTGTTGGGTATATCATTTAA 1201
Db
```

## RESULT 4

US-08-105-483-215  
Sequence 215, Application US/08105483

Patent No. 5494807  
GENERAL INFORMATION:  
APPLICANT: Paolietti, Enzo  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
TITLE OF INVENTION: STRAIN  
NUMBER OF SEQUENCES: 462  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESS: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

```
/?
/? COMPUTER READABLE FORM:
/? MEDIUM TYPE: Floppy disk
/? COMPUTER: IBM PC compatible
/? OPERATING SYSTEM: PC-DOS/MS-DOS
/? SOFTWARE: Patent in Release #1.0, Version #1.25
/? CURRENT APPLICATION DATA:
/? APPLICATION NUMBER: US/08/105,483
/? FILING DATE: 12-AUG-1993
/? CLASSIFICATION: 424
/? PRIOR APPLICATION DATA:
/? APPLICATION NUMBER: US 07/847,951
/? FILING DATE: 06-MAR-1992
/? ATTORNEY/AGENT INFORMATION:
/? NAME: Frommer, William S.
/? REGISTRATION NUMBER: 25,506
/? REFERENCE/DOCKET NUMBER: 454310-2400
/? TELEPHONE: (212) 840-3333
/? TELEFAX: (212) 840-0712
/? INFORMATION FOR SEQ ID NO: 215:
/? SEQUENCE CHARACTERISTICS:
/? LENGTH: 1285 base pairs
/? TYPE: nucleic acid
/? STRANDEDNESS: single
/? TOPOLOGY: linear
/?
/? US-08-105-483-215
```

Query Match 94.6%; Score 1117.2; DB 1; Length 1285;

Best Local Similarity 97.2%; Pred. No. 0; Mismatches 33; Indels 0; Gaps 0;

Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

1 ATGGGCGAGATCTTTTCCACAGCAATCTCTGGATTTCTTCCGACACACAGTTGGAT 60

116 ATGGGCGAGATCTTTTCCACAGCAATCTCTGGATTTCTTCCGACACACAGTTGGAT 125

61 CAGCCTTCAGACCAACACCAATTCAGATTGGACTTCAATCCCAAGACACC 120

176 CAGCCTTCAGACCAACACCAATTCAGATTGGACTTCAATCCCAAGACACC 235

121 TTGGCGAGCGCCAAACAGGATGAGAGCTGAGCATTTGGAGTTCCACCCACCGAC 180

236 TTGGCGAGCGCCAAACAGGATGAGAGCTGAGCATTTGGAGTTCCACCCACCGAC 235

181 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACCTTGGCAGCAAT 240

296 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACCTTGGCAGCAAT 355

241 CGGCTCTCTGCTTCACCAATGCGCAGTCAAGAGGACGCTAACCCGCTGTCCAAT 300

356 CGGCTCTCTGCTTCACCAATGCGCAGTCAAGAGGACGCTAACCCGCTGTCCAAT 415

301 TTGAGAAACATCATCTCTCAAGCATGAGTGAATCCCAACCTTCCACCAACTCTG 360

416 TTGAGAAACATCATCTCTCAAGCATGAGTGAATCCCAACCTTCCACCAACTCTG 475

361 CAAGATCCAGAGTGAAGTCTGTATTTCCCTGCTGTGTGCTCCAGTTGAGAAACAGTA 420

476 CAAGATCCAGAGTGAAGTCTGTATTTCCCTGCTGTGTGCTCCAGTTGAGAAACAGTA 535

421 AACCCGTTCCGACTCTGTCTCTCCATATGCTCAATCTTCTGAGAGATTGGGACCT 480

536 AACCCGTTCCGACTCTGTCTCTCCATATGCTCAATCTTCTGAGAGATTGGGACCT 595

481 GCGCGAACAATGAGAAACATCAGATCAGATTTCTAGAGACCCCTGCTCGTTTACAGCG 540

596 GCGCGAACAATGAGAAACATCAGATCAGATTTCTAGAGACCCCTGCTCGTTTACAGCG 655

541 GGGTTTTCTTTGTTGACAAATCTTCAATAACCGAGTCTAGACTGCTGGTGACT 600

656 GGGTTTTCTTTGTTGACAAATCTTCAATAACCGAGTCTAGACTGCTGGTGACT 715

601 TCTCTCAATTTTCTAGGGGGAACACTACCGTGTCTTGGCCAAAATTCAGAGTCCCAAC 660

Db 716 TCTCTCAATTTTCTAGGGGAACTACCGTGTCTTTGGCAAAATTCGACGTCCCAACC 775  
 Qy 661 TCCATCACTACCAACCTCTGTCTCTCAACTTGTCTGTATGCGTGTATGTCGTG 720  
 Db 776 TCCATCACTACCAACCTCTGTCTCTCAACTTGTCTGTATGCGTGTATGTCGTG 835  
 Qy 721 CGGCGTTTATCATCTTCTCTTTCATCTGTCTGTATGCGTGTATGTCGTG 780  
 Db 836 CGGCGTTTATCATCTTCTCTTTCATCTGTCTGTATGCGTGTATGTCGTG 895  
 Qy 781 CTGACATCAAGATATGTTGCCGTTTGTCTCTTATTCAGATCTTCAACACACAC 840  
 Db 896 CTGACATCAAGATATGTTGCCGTTTGTCTCTTATTCAGATCTTCAACACACAC 955  
 Qy 841 ACGGACCATGACAGCTGTACAGCTCTGTCTCAAGAACTCTATGATCCCTCTGT 900  
 Db 956 ACGGACCATGACAGCTGTACAGCTCTGTCTCAAGAACTCTATGATCCCTCTGT 1015  
 Qy 901 TCGTGTACAAAACCTTGGAGTGAACCTGACCTGTATTCACATCCATCATCTGGGCT 960  
 Db 1016 TCGTGTACAAAACCTTGGAGTGAACCTGACCTGTATTCACATCCATCATCTGGGCT 1075  
 Qy 961 TTGCGAAATTCCTATGAGAGTGGGCTCTACGCCCTTCTCTGGCTCAGTTACTAGTG 1020  
 Db 1076 TTGCGAAATTCCTATGAGAGTGGGCTCTACGCCCTTCTCTGGCTCAGTTACTAGTG 1135  
 Qy 1021 CCATTTGTTCAGTGTGTGTAGGGCTTTCGCCCATCTGTGCTTCACTATGATGAGATG 1080  
 Db 1136 CCATTTGTTCAGTGTGTGTGTAGGGCTTTCGCCCATCTGTGCTTCACTATGATGAGATG 1195  
 Qy 1081 ATGTTGTATGAGAGTGGGCTCTACAGCTCTGTACACATCTGTGAGTCCCTTTTACCGCTGTTACCA 1140  
 Db 1196 ATGTTGTATGAGAGTGGGCTCTACAGCTCTGTACAGATCTGTGAGTCCCTTTTACCGCTGTTACCA 1255  
 Qy 1141 ATTTTCTTTTGTCTTGTGGGTATACATTAA 1170  
 Db 1256 ATTTTCTTTTGTCTTGTGGGTATACATTAA 1285

## RESULT 5

US-08-709-209-215  
 ; Sequence 215, Application US/08709209  
 ; Patent No. 5762938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paoletti, Enzo  
 ; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
 ; TITLE OF INVENTION: STRAIN  
 ; NUMBER OF SEQUENCES: 462  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/709,209  
 ; FILING DATE: 21-AUG-1996  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/105,483  
 ; FILING DATE: 12-AUG-1993  
 ; APPLICATION NUMBER: US 07/847,951  
 ; FILING DATE: 06-MAR-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frommer, William S.  
 ; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454310-2400  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; INFORMATION FOR SEQ ID NO: 215:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1285 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-709-209-215  
 Query Match 94.6%; Score 1117.2; DB 1; Length 1285;  
 Best Local Similarity 97.2%; Pred. No. 0;  
 Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ATGGGGCAGATCTTTCACACGACATCTCTGGGATCTTTCGACACCAAGTTGGAT 60  
 Db 116 ATGGGGCAGAACTCTTTCACACGACATCTCTGGGATCTTTCGACACCAAGTTGGAT 175  
 Qy 61 CCAGCTTCAGAGCAACCAACCAATCCAGATGGGACTTCAATCCCAACGAAGACACC 120  
 Db 176 CCAGCTTCAGAGCAACCAACCAATCCAGATGGGACTTCAATCCCAACGAAGACACC 235  
 Qy 121 TGGCCAGACGCGCAACAGATGAGAGCTGAGCATTCGAGCTGGGGTTGACCCACGAC 180  
 Db 226 TGGCCAGACGCGCAACAGATGAGAGCTGAGCATTCGAGCTGGGGTTGACCCACGAC 295  
 Qy 181 GGAGGCTTTTGGGGTGAAGCTTCAGGCTCAGGGCATPAACAAACCTTGCAGCAAT 240  
 Db 296 GGAGGCTTTTGGGGTGAAGCTTCAGGCTCAGGGCATPAACAAACCTTGCAGCAAT 355  
 Qy 241 CGGCTCTCTGCTTCCACCAATTCGACAGAGAGGACCTTACCCGCTGTCTCACT 300  
 Db 356 CGGCTCTCTGCTTCCACCAATTCGACAGAGAGGACCTTACCCGCTGTCTCACT 415  
 Qy 301 TTGAGAAACACTCATCTCAAGCCATGACATGAGAGTCAACCACTTCCACCAACTGTG 360  
 Db 416 TTGAGAAACACTCATCTCAAGCCATGACATGAGAGTCAACCACTTCCACCAACTGTG 475  
 Qy 361 CAAGATCCCAAGTGAAGTGTGATTTCCCTGCTGTGTGCTCAAGTTACAGAAACAGTA 420  
 Db 476 CAAGATCCCAAGTGAAGTGTGATTTCCCTGCTGTGTGCTCAAGTTACAGAAACAGTA 535  
 Qy 421 AACCTGTTCGACATCTCTCTCCCATTCGTAATCTTCTGAGATGGGGACCT 480  
 Db 536 AACCTGTTCGACATCTCTCTCCCATTCGTAATCTTCTGAGATGGGGACCT 595  
 Qy 481 GCGCGAATGAGAAATCATCATCAGATTCCTAGAGCCCTCTGCTGTTACAGGCG 540  
 Db 596 GCGCGAATGAGAAATCATCATCAGATTCCTAGAGCCCTCTGCTGTTACAGGCG 655  
 Qy 541 GGGTTTTTCTGTGACAAAGATCTCTCAATACCGAGAGTCTGAGTCTGTGAGT 600  
 Db 656 GGGTTTTTCTGTGACAAAGATCTCTCAATACCGAGAGTCTGAGTCTGTGAGT 715  
 Qy 601 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTCTGTGGCAAAATTCGACAGTCCCAACC 660  
 Db 716 TCTCTCAATTTTCTAGGGGGAATCTACCGTGTCTGTGGCAAAATTCGACAGTCCCAACC 775  
 Qy 661 TCCATCACTACCAACCTCTGTCTCTCAACTTGTCTGTATGCGTGTATGTCGTG 720  
 Db 776 TCCATCACTACCAACCTCTGTCTCTCAACTTGTCTGTATGCGTGTATGTCGTG 835  
 Qy 721 CGGCGTTTATCATCTTCTCTTTCATCTGTCTGTATGCGTGTATGTCGTG 780  
 Db 836 CGGCGTTTATCATCTTCTCTTTCATCTGTCTGTATGCGTGTATGTCGTG 895  
 Qy 781 CTGACATCAAGATATGTTGCCGTTTGTCTCTTATTCAGATCTTCAACACACAC 840  
 Db 896 CTGACATCAAGATATGTTGCCGTTTGTCTCTTATTCAGATCTTCAACACACAC 955  
 Qy 841 ACGGACCATGACAGCTGTACAGCTCTGTCTCAAGAACTCTATGATCCCTCTGT 900

```

Db      956  ACCGGACCATGCCCAACCTCATGACTGCTCAAGAACTCTATGATCCCTCTGT
Qy      901  TGGGTACAAACCTTGAGATGGAACCTGATTCCTCATCCATCATCTGGGCT
Db      1016  TGGGTACAAACCTTGAGATGGAACCTGATTCCTCATCCATCATCTGGGCT
Qy      961  TTGGAAATTCCTATGAGAGTGGGCTCAAGCCGTTCTCTGGCTCACTTACTAGT
Db      1076  TTGGAAATTCCTATGAGAGTGGGCTCAAGCCGTTCTCTGGCTCACTTACTAGT
Qy      1021  CCATTTGTCAAGTGTCTGAGAGCTTCCGCCCTGTTGGCTTCACTTATGAGATG
Db      1136  CCAATTGTCAAGTGTCTGAGAGCTTCCGCCCTGTTGGCTTCACTTATGAGATG
Qy      1081  ATGTGTACTGGGGGCCAAGCTGTATACCATCTGAGTCCCTTTTACCGCTGTACCA
Db      1196  ATGTGTACTGGGGGCCAAGCTGTATACCATCTGAGTCCCTTTTACCGCTGTACCA
Qy      1141  ATTTCTTTTGTCTTTGGGTATACATTTAA
Db      1256  ATTTCTTTTGTCTTTGGGTATACATTTAA

```

# RESULT 6

```

US-08-458-101-215
; Sequence 215, Application US/08458101
; Patent No. 5766599
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Perkins, Marion E.
; APPLICANT: Taylor, Jill
; APPLICANT: Tartaglia, James
; APPLICANT: No. 5766599, Elizabeth K.
; APPLICANT: Riviere, Michel
; APPLICANT: de Taisne, Charles
; APPLICANT: Limbach, Keith J.
; APPLICANT: Johnson, Gerard P.
; APPLICANT: Pincus, Steven E.
; APPLICANT: Cox, William I.
; APPLICANT: Audonnet, Jean-Christophe Francis
; APPLICANT: Gettig, Russell Robert
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,101
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 215:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-458-101-215
Query Match      94.6%; Score 1117.2; DB 1; Length 1285;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy      1  ATGGGAGAGATCTTTCACACAGCAATCTCTGGAGATCTTTCACAGCACAGTTGAT 60
Db      116  ATGGGAGAGATCTTTCACACAGCAATCTCTGGAGATCTTTCACAGCACAGTTGAT 175
Qy      61  CCAAGCTTCAGACCAACCAACCAATCCAGATTGGAGCTTCAATCCCAAGACACC 120
Db      176  CCAAGCTTCAGACCAACCAACCAATCCAGATTGGAGCTTCAATCCCAAGACACC 235
Qy      121  TGGCCAGAGCCCAACCAAGTAGAGGTGAGCAATGGAGCTGGGGTTCAACCCAGCCAC 180
Db      236  TGGCCAGAGCCCAACCAAGTAGAGGTGAGCAATGGAGCTGGGGTTCAACCCAGCCAC 295
Qy      181  GAGAGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAAACCTTGGCCAGCAAT 240
Db      296  GAGAGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAAACCTTGGCCAGCAAT 355
Qy      241  CGGCTCTCTGCTTCCACCAATCGCCAGTACAGAAAGCCTAACCCGCTGTCTCACT 300
Db      356  CGGCTCTCTGCTTCCACCAATCGCCAGTACAGAAAGCCTAACCCGCTGTCTCACT 415
Qy      301  TTGAGAAACATCATCTCTCAAGCATGAGTGAATCCCAACCTTCCACCAACTCTG 360
Db      416  TTGAGAAACATCATCTCTCAAGCATGAGTGAATCCCAACCTTCCACCAACTCTG 475
Qy      361  CAAGATCCAGAGTGAAGGCTGTATATTTCCCTGTGTGCTCAGTTACAGAAACATG 420
Db      476  CAAGATCCAGAGTGAAGGCTGTATATTTCCCTGTGTGCTCAGTTACAGAAACATG 535
Qy      421  AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTCTGAGAGATTGGGACCT 480
Db      536  AACCTGTTCGACTACTGTCTCTCCATATCGTCAATCTTCTGAGAGATTGGGACCT 595
Qy      481  GCGGGGAACATGAGAAACATCATCAGAGTTCTTGGGACCCCTGCTGGTTACAGGG 540
Db      596  GCGGTGAACATGAGAAACATCATCAGAGTTCTTGGGACCCCTGCTGGTTACAGGG 655
Qy      541  GGGTTTTTCTTTTGAAGAATCTCACAATCCGAGAGTCTAGACTGTGTGAGCT 600
Db      656  GGGTTTTTCTTTTGAAGAATCTCACAATCCGAGAGTCTAGACTGTGTGAGCT 715
Qy      601  TCTCTCAATTTCTAGGGGAATCACTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 660
Db      716  TCTCTCAATTTCTAGGGGAATCACTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 775
Qy      661  TCCAAATCACTACCAACCTCTGCTCCCAACTGTCTGGTATGCTGGAGATGTGCTG 720
Db      776  TCCAAATCACTACCAACCTCTGCTCCCAACTGTCTGGTATGCTGGAGATGTGCTG 835
Qy      721  CGGCGTTTATCATCTTCTCTTCACTCTGCTGCTATGCTCACTCTTGTGTTCTT 780
Db      836  CGGCGTTTATCATCTTCTCTTCACTCTGCTGCTATGCTCACTCTTGTGTTCTT 895
Qy      781  CTGAGACTATCAAGATATGTTGCCGTTTGTCTCTTAATTCAGAGATCTTCAACACAC 840
Db      896  CTGAGACTATCAAGATATGTTGCCGTTTGTCTCTTAATTCAGAGATCTTCAACACAC 955
Qy      841  ACCGGAACATGACAGCTGCAAGCTCTGCTCAAGGAACCTTATGATCCCTCTGT 900
Db      956  ACCGGAACATGACAGCTGCAAGCTCTGCTCAAGGAACCTTATGATCCCTCTGT 1015
Qy      901  TGGGTACAAACCTTGAGATGGAACCTGATTCCTCATCCATCATCTGGGCT 960
Db      1016  TGGGTACAAACCTTGAGATGGAACCTGATTCCTCATCCATCATCTGGGCT 1075
Qy      961  TTGGAAATTCCTATGAGAGTGGGCTCAAGCCGTTTCTCTGGCTCACTTACTAGT 1020

```



Db 1076 TTCGAAATTCCTATAGGAGTGAGGCTCAGCCCGTTCTCCGTGAGTCTAGTATAGTG 1135  
Qy 1021 CCATTGTCAGTGGTTCGTAGAGCTTTCCCCACCTTTGGCTTCACTTATATGAGTG 1080  
Db 1136 CCATTGTTCAGTGGTTCGTAGAGCTTTCCCCACCTTTGGCTTCACTTATATGAGTG 1195  
Qy 1081 ATGTTGTAAGGAGGCGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTACCA 1140  
Db 1196 ATGTGTATTTGGGGGCGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTACCA 1255  
Qy 1141 ATTTCTTTTGTCTTTGGGTATACATTAA 1170  
Db 1256 ATTTCTTTTGTCTTTGGGTATACATTAA 1285  
RESULT 7  
US-08-480-173A-42  
; Sequence 42, Application US/08480173A  
; Patent No. 6072049  
; GENERAL INFORMATION:  
; APPLICANT: Thoma, Hans A  
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Popovich & Miles, P.A.  
; STREET: 80 S. 8th Street, Suite 1902  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,173A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Popovich, Thomas E  
; REGISTRATION NUMBER: 30,099  
; REFERENCE/DOCKET NUMBER: MED1003USD4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-334-8991  
; TELEFAX: 612-334-8994  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2348 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 12..14  
; OTHER INFORMATION: /note= "S1 start codon"  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 336..338  
; OTHER INFORMATION: /note= "S2 start codon"  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 508..510  
; OTHER INFORMATION: /note= "S start codon"  
US-08-480-173A-42  
Query Match 92.1%; Score 1088.2; DB 3; Length 2348;  
Best Local Similarity 96.5%; Pred. No. 0;  
Matches 1147; Conservative 0; Mismatches 33; Indels 9; Gaps 3;  
Qy 1 ATGGGCGAAGATCTTCCACGACGACATCTCTGAGATCTTCCGACGACGAGTTGAT 60

Db 12 ATGGGCGAAGATCTTCCACGACGACATCTCTGAGATCTTCCGACGACGAGTTGAT 71  
Qy 61 CCAAGCTTCAGAGCAAAACCAACAATCCAGATTTGGACTTAAATCCCAAGAGACACC 120  
Db 72 CCAAGCTTCAGAGCAAAACCAACAATCCAGATTTGGACTTAAATCCCAAGAGACACC 131  
Qy 121 TGGCCAGAGCGCCAAAGGTAGAGCTGAGACATTGGAGCTGGGGTTACACCCACGAC 180  
Db 132 TGGCCAGAGCGCCAAAGGTAGAGCTGAGACATTGGAGCTGGGGTTACACCCACGAC 191  
Qy 181 GGAGGCGTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACAACCTTGGCAGCAAT 240  
Db 192 GGAGGCGTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACAACCTTGGCAGCAAT 251  
Qy 241 CGGCTCTCTGCTTCCACCAATGCGCATGAGAGAGGACCTTACCCCGCTGTCTCAGCT 300  
Db 252 CGGCTCTCTGCTTCCACCAATGCGCATGAGAGAGGACCTTACCCCGCTGTCTCAGCT 311  
Qy 301 TTGAGAAACATCATCTCAAGCCATGAGTGAAGTCCACAACCTTCCACCAACTCTG 360  
Db 312 TTGAGAAACATCATCTCAAGCCATGAGTGAAGTCCACAACCTTCCACCAACTCTG 371  
Qy 361 CAAGATCCAGAGTGAAGTCTGTATTTCCCTGCTGTGCT-----CCAGTTTCAG 413  
Db 372 CAAGATCCAGAGTGAAGTCTGTATTTCCCTGCTGTGCTCCAGTTCCACGTTTCAG 431  
Qy 414 AACAGTAAACCTGTTCCGACTACTGTCTTCCCATATGCTCATCTTTCGAGATTGG 473  
Db 432 AACAGTAAACCTGTTCCGACTACTGTCTTCCCATATGCTCATCTTTCGAGATTGG 491  
Qy 474 GGACCTCGGCGGAGATGAGAGACATCATCAGATTTCCAGAGACCCCTGCTGTGTT 533  
Db 492 GGACCTCGGCGGAGATGAGAGACATCATCAGATTTCCAGAGACCCCTGCTGTGTT 551  
Qy 534 ACAGGCGGGGTTTTTTCTTTGACAAAGATCTCAAAATACCGCAGAGTCTAGACTCGT 593  
Db 552 ACAGGCGGGGTTTTTTCTTTGACAAAGATCTCAAAATACCGCAGAGTCTAGAGT 610  
Qy 594 GTGACTTCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGCAATC 653  
Db 611 GTGACTTCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGCAATC 670  
Qy 654 CCAACCTCCATCACTACCAACCTCCGTCTCCCACTGTCTGTTATTCGCTGAT 713  
Db 671 CCAACCTCCATCACTACCAACCTCCGTCTCCCACTGTCTGTTATTCGCTGAT 730  
Qy 714 GTGTCTGGGGGTTTTTATCATCTTCTCTTCACTCGTGTATGCTCATCTTCTGTT 773  
Db 731 GTGTCTGGGGGTTTTTATCATCTTCTCTTCACTCGTGTATGCTCATCTTCTGTT 790  
Qy 774 GGTCTTTCGACATACAAAGTATGTGCCGTTTTGCTCTTAATTCAGAGATCTTCAAC 833  
Db 791 GGTCTTTCGACATACAAAGTATGTGCCGTTTTGCTCTTAATTCAGAGATCTTCAAC 850  
Qy 834 CACGACGACGAGACCATGACAGCCTGCAGACCTCTCTCAAGGAACCTCATATATCC 893  
Db 851 AACGACGACGAGACCATGACAGCCTGCAGACCTCTCTCAAGGAACCTCATATATCC 910  
Qy 894 CTCCTGTGCTGTACAAACCTTCGATGAGAACTGACACTGTATTTCCATCCATCATC 953  
Db 911 CTCCTGTGCTGTACAAACCTTCGATGAGAACTGACACTGTATTTCCATCCATCATC 970  
Qy 954 CTGGGCTTTCGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTT 1013  
Db 971 CTGGGCTTTCGAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCAGTTT 1030  
Qy 1014 ACTAGTGCATTTGTTCAAGTGTCTGAGGCTTTCGCCCATGTTTGGCTTCAAGTAT 1073  
Db 1031 ACTAGTGCATTTGTTCAAGTGTCTGAGGCTTTCGCCCATGTTTGGCTTCAAGTAT 1090  
Qy 1074 ATGATGATGTTGTACTGGGGGCGAAGTCTG-TACACCATCTTGAGTCCCTTTTACCGC 1132



```
Sequence 5, Application US/08854531
Patent No. 6025341
GENERAL INFORMATION:
APPLICANT: Wands, Jack
APPLICANT: Tokushige, Katsutoshi
APPLICANT: Wakita, Takaji
TITLE OF INVENTION: CHIMERIC HEPATITIS B/HEPATITIS C VIRUS VACCINE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSER: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6025341rirs
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,531
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: APOI-0214
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1200 base pairs
TYPE: nucleic acid
STANDARDS: double
TOPOLOGY: linear
US-08-854-531-5

Query Match      84.7%; Score 1000.6; DB 3; Length 1200;
Best Local Similarity 91.1%; Pred. No. 1,8e-294;
Matches 1063; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
```

```
Db 454 AACCTGTTCCGACTACTGTCTCACCATATGCTCAATCTCTCGAGAGACTGGGACCT 513
Qy 481 GCGCGGACATGAGAGACATCAATCAGGATTCCTAGGACCCCTGCTGTTACAGGG 540
Db 514 GCACGAGACATGAGAGACATCAATCAGGATTCCTAGGACCCCTGCTGTTACAGGG 573
Qy 541 GGGTTTTCTTTGTCAGAGAAATCTCAGATACGACAGCTGAGCTGAGTGGAGT 600
Db 574 GGGTTTTCTTTGTCAGAGAAATCTCAGATACGACAGCTGAGCTGAGTGGAGT 633
Qy 601 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 660
Db 634 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAAC 693
Qy 661 TCCATCATCTCAACACCTCTGCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 694 TCCATCATCTCAACACCTCTGCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTG 753
Qy 721 CCGCGTTTATCATCTCTCTTCAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 754 CCGCGTTTATCATCTCTCTTCAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
Qy 781 CTGACTATCAAGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 814 CTGACTATCAAGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
Qy 841 AGGAGACATGAGAGACCTGACAGCTCTGCTCAAGGAACTCTATGATCCCTCTCT 900
Db 874 AGGAGACATGAGAGACCTGACAGCTCTGCTCAAGGAACTCTATGATCCCTCTCT 933
Qy 901 TCGTGTACAAACCTTCGAGTGAAGAACTGACCTGATTCCTCATCCATCATCTGAGCT 960
Db 934 TCGTGTACAAACCTTCGAGTGAAGAACTGACCTGATTCCTCATCCATCATCTGAGCT 993
Qy 961 TTGGAGAAATTCCTATGAGAGTGGGCTTCAGCCCTGCTGCTGCTGCTGCTGCTG 1020
Db 994 TTGGAGAAATTCCTATGAGAGTGGGCTTCAGCCCTGCTGCTGCTGCTGCTGCTG 1053
Qy 1021 CCATTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1054 CCATTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1113
Qy 1081 ATGTGTACTGGGAGCAAGCTGTACACCATCTGAGTCCCTTTTACCGCTTTACCA 1140
Db 1114 ATGTGTACTGGGAGCAAGCTGTACACCATCTGAGTCCCTTTTACCGCTTTACCA 1173
Qy 1141 ATTTCTTTTGTCTTGGGTATACAT 1167
Db 1174 ATTTCTTTTGTCTTGGGTATACAT 1200

RESULT 10
PCT-US95-13552-5
Sequence 5, Application PC/TUS9513552
GENERAL INFORMATION:
APPLICANT: Wands, Jack
APPLICANT: Tokushige, Katsutoshi
APPLICANT: Wakita, Takaji
APPLICANT: Pachuk, Catherine J.
APPLICANT: Zurawski, Jr., Vincent R.
APPLICANT: Coney, Leslie R.
TITLE OF INVENTION: HEPATITIS VIRUS VACCINES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSER: Woodcock, Washburn, Kurtz, Mackiewicz &
ADDRESSER: Norris
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13552  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/318,248  
FILING DATE: 05-OCT-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/467,859  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: APOL-0238  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEO ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1200 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
PCT-US95-13552-5

Query Match 84.7%; Score 1000.6; DB 5; Length 1200;  
Best Local Similarity 91.1%; Pred. No. 1.8e-294;  
Matches 1063; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 1 ATGGGGAGATCTTTCCACGAGATCTCTGGGATTTCCCGACCAAGATGGAT 60  
DB 34 ATGGGGAGATCTTTCTGTTCCCAATCTCTGGGATTTCCCGATTCACCAATGGAC 93  
QY 61 CCAAGCTTCAGAGCAACCAACCAATTCAGATTGGGACTTCAATCCCAACAGACACC 120  
DB 94 CTGCGCTTCGAGGCAACTCAACATTCAGATTGGGACTTCAACCCCAACAGATCA 153  
QY 121 TGCCAGACGCCAACAAGTAGAGCTGAGCATTCGGATCGGGTTTACCCACCGCAC 180  
DB 154 TGCCAGAGGAATCAAGTAGAGCGGAGACTTCGGGTCAGGGTTTACCCACACAC 213  
QY 181 GGAGGCTTTGGGGTGAAGCCCTCAGGCTCAGGGCATACACAACCTTCCAGCAAT 240  
DB 214 GGGGCTTTGGGGTGAAGCCCTCAGGCTCAGGGCATATTGACACAAGTCCAGACGG 273  
QY 241 CCGCTCTCTGCTTCACCAATCGCAGTCAGAGAGGAGCTTACCCGCTGTCTCACT 300  
DB 274 CTCTCTCTGCTTCACCAATCGCAGTCAGAGAGGAGCTTACCCGCTGTCTCACT 333  
QY 301 TTGAGAAACATCTCTCAAGCCATGAGTGAATTCACAACCTTTCACCAAACTCTG 360  
DB 334 CTAGAGAGAGTCACTCTCAGGCGCATGAGTGAATTCACAACATTCACCAAGCTCTG 393  
QY 361 CAATATCCCAAGAGAGTCTATTTCCCTGCTGTGCTCAAGTTCAGAGAACTGA 420  
DB 394 CTAGATCCCAAGAGAGTCTATTTCCCTGCTGTGCTCAAGTTCAGAGAACTGA 453  
QY 421 AACCTGTTCGATCTCTCTCCATATGTCATCTCTGAGGATGGGAGACCT 480  
DB 454 AACCTGTTCGATCTCTCTCCATATGTCATCTCTGAGGATGGGAGACCT 513  
QY 481 GCGGGAACATGAGAACATCATCAGATTCCTAGAGACCTGCTGTGTTCAGAGCG 540  
DB 514 GCACCGAAGATGAGAGACAAATCAGGATTCCTAGAGACCTGCTGTGTTCAGAGCG 573  
QY 541 GGGTTTTTCTGTGACAAAGATCTCTCAATACCGCAGAGTCTAGACTGTGTGACT 600  
DB 574 GGGTTTTTCTGTGACAAAGATCTCTCAATACCGCAGAGTCTAGACTGTGTGACT 633

QY 601 TCTCTCAATTTCTAGGGGAACTACCGTGTCTTGGCCAAATTCGAGTCCCAACC 660  
DB 634 TCTCTCAATTTCTAGGGGAGACCAAGTGTCTGCGCAAAATTCGAGTCCCAACC 693  
QY 661 TCCATATCTACCAACCTCTCTCTCTCAACTTGTCTGTATATGCTGTGATGTCG 720  
DB 694 TCCATATCTACCAACCTCTCTCTCTCAACTTGTCTGTATATGCTGTGATGTCG 753  
QY 721 CCGGCTTTTATCATCT 780  
DB 754 CCGGCTTTTATCATCT 813  
QY 781 CTGAGATATCAAGATATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACACAGC 840  
DB 814 CTGAGATATCAAGATATGTTGCCGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 873  
QY 841 ACGGAGCATGACAGAGCTCTGACAGCTCTGCTCAAGAACTCTATATATCTCTCTCT 900  
DB 874 ACGGAGCATGACAGAGCTCTGACAGATCTCTGCTCAAGAACTCTATATGTTTCCCTCT 933  
QY 901 TGCTGTACAAACCTTCGATGAGAACTGACCTGTATTCATCCATCATCTCTGAGCT 960  
DB 934 TGCTGTACAAACCTTCGATGAGAACTGACCTGTATTCATCCATCATCTCTGAGCT 993  
QY 961 TTGGAAAATTCCTATGAGAGTGGGCTCAGCCGTTTCTCTGCTCACTTACTAGTG 1020  
DB 994 TTGCAAGATTCCTATGAGAGTGGGCTCAGCTCGTTTCTCTGCTCACTTACTAGTG 1053  
QY 1021 CCATTGTTCAGTGTCTGAGGCTTTCCCGCACTTTGGCTTCAGTATATGATG 1080  
DB 1054 CCAATTGTTCAGTGTCTGAGGCTTTCCCGCACTTTGGCTTCAGTATATGATG 1113  
QY 1081 ATGTGTATGAGGAGCAAGTCTGTACACCATCTTGTAGTCCCTTTTACCGCTGTAACA 1140  
DB 1114 ATGTGTATGAGGAGCAAGTCTGTACACCATCTTGTAGTCCCTTTTACCGCTGTAACA 1173  
QY 1141 ATTTCTTTTGTCTTTGGTATACAT 1167  
DB 1174 ATTTCTTTTGTCTTTGGTATACAT 1200

RESULT 11  
US-08-715-808-2  
Sequence 2, Application US/08715808  
Patent No. 5981274  
GENERAL INFORMATION:  
APPLICANT: Tyrrell, D. Lorne J.  
APPLICANT: Chaisomchit, Sumonta  
TITLE OF INVENTION: Recombinant Hepatitis Virus Vectors  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medien & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/715,808  
FILING DATE: 18-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ingolia, Diane E.  
REGISTRATION NUMBER: 40,027  
REFERENCE/DOCKET NUMBER: CHANG-02441  
TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9325 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "DNA"
US-08-715-808-2

```

Query Match	84.7%	Score 1000.2	DB 2	Length 9325
Best Local Similarity	90.4%	Pred. No. 6.5e-294		
Matches 1068	Conservative	0	Mismatches 113	Indels 0
			Gaps	0
Qy	1	ATGGGGCAGAAATCTTTCCAGCAGCATCTCTGGGATTTCTTCCGACACCCAGTGGAT	60	
Db	3135	ATGGGGAAGAAATCTTTCTGTTTCCAAATCCTCTGGGATTTCTTCCGATATCATGATTGGAC	3194	
Qy	61	CCAGCTTTCAGAGCAACACCAACATTCAGATTGGAGCTTCAATCCCAAGAGACAC	120	
Db	3195	CCTGATTTGGGAGCCAACTCAAAACATCCAGATTGGGACTTCAACCCCGTCAGAGCAGC	325	
Qy	121	TGGCCAGACGCCAACAAGTAGSACTGAGATTCGACTCGGGTTTCAACCCACCGAC	180	
Db	3255	TGGCCAGAGCCCAACCAAGTAGAAGTGGAGATTTCGGGCCAAGGCTCAACCCCTCCAC	331	
Qy	181	GGAGGCTTTTGGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACAACCTTGGCAGAAAT	240	
Db	3315	GGCGGTAATTTTGGGGTGAAGCCCTCAGGCTCAGGGCATTTTGAACCAACAGTGTCAACAT	337	
Qy	241	CGGCTCTGCTTCCACCAATCGCAGTTCAGGAGGACAGCCCTAACCCGCTGCTCCACT	300	
Db	3375	CCTCTCTGCTCTCACCACATCGGAGTTCAGGAGGACAGCCCTAATCCCATCTCTCCACT	343	
Qy	301	TTGAGAAACATCATCTCTCAGGCTCAGTGGAACTCACAACATTTTCCACCAACTCTG	360	
Db	3435	CTAAGAGACAGTCAATCTCAGGCTCAGTGGAAATTCACATGCTTCCACCAACTCTG	349	
Qy	361	CAAGATCCCAAGTAGAGAGTGTATTTCCCTGCTGGGTGCTCCAGTTCAGAAACAGTA	420	
Db	3495	CAGATCCCAAGTAGAGGAGTGTATTTCTTCTGCTGGGTGCTCCAGTTCAGAAACAGTA	355	
Qy	421	AACCTGTTCGAGTACTGTCTCTCCATATGTCATCTTTCGAGATTGGGAGCCCT	480	
Db	3555	AACCTGTCTCGAATATTGCTCTCAATCTGTCATCTCGCGAGGAGCTGGGAGCCCT	361	
Qy	481	GCGCGAAACATGAGAACATCATCATCAGGATCTCTAGAGACCCCTGCTGTTACAAGCG	540	
Db	3615	GTCAGGAACATGAGAACATCATCATCAGGATCTCTAGAGACCCCTGCTGTTACAAGCG	367	
Qy	541	GGGTTTTCTGTTGACAAAGATCTCTCAACAAACCGAGAGTCTAGACTCTCGGAGGACT	600	
Db	3675	GGGTTTTCTGTTGACAAAGATCTCTCAACAAACCGAGAGTCTAGACTCTCGGAGGACT	373	
Qy	601	TCTCTCAATTTTCTAGGGGGAATCTACCGTGTCTTGGCCAAATTCGAGTCCCAAC	660	
Db	3735	TCTCTCAATTTTCTAGGGGGAATCTCCCGTGTCTTGGCCAAATTCGAGTCCCAAC	379	
Qy	661	TCCATCACTCAACCACTCTGTCTTCCACTTGTCTGTATTGCTGAGTGTCTG	720	
Db	3795	TCCATCACTCAACCACTCTGTCTTCCACTTGTCTGTATTGCTGAGTGTCTG	385	
Qy	721	CGGGTTTTATCATTTCTCTTCAATCCGCTGCTATAGCTCATCTTCTTGTGGTCTT	780	
Db	3855	CGGGTTTTATCATTTCTCTTCAATCCGCTGCTATAGCTCATCTTCTTATTTGGTCTT	391	
Qy	781	CTGCACTATCAAGTATGTGCGGTTGTCTCTTAATCCAGATCTTCAACCAACAGC	840	
Db	3915	CTGCACTATCAAGTATGTGCGGTTGTCTCTTAATCCAGATCAACCAACCAAGT	397	
Qy	841	ACGGGACATGACAGCCTGACGACTCTGTCTCAAGAACTCTATGTTATCTCTGT	900	

Db	3975	ACGGGACCATGCAAAACCTGCACACTCTCGTCTCAAGCAACTGTATGTTTCCCTCANGT	4034
Qy	901	TGCTGTACAAACCTTCGATGGAACCTGCACCTGTATTCACATCCCATATCTTGAGCT	960
Db	4035	TGCTGTACAAACCTTACGAGATGGAATTTGCACCTGTATTCACATCCCATATCTTGAGCT	4094
Qy	961	TTGGGAAATTTCCATATGGGAGTGGGGCTCAGCCCGTTTTCCTGGGCTCAGTTTACTAGTG	1020
Db	4095	TTCCAAAAATTCCTATATGGAGTGGGGCTCAGCTCGTTTCTCTGGCTCAGTTTACTAGTG	4154
Qy	1021	CCATTGTTCAGTGGTTCGATAGGGCTTTCCGCCACTGTTTGGCTTCAGTTATATGATG	1080
Db	4155	CCATTGTTCAGTGGTTCGATAGGGCTTTCCGCCACTGTTTGGCTTCAGTTATATGATG	4214
Qy	1081	ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTAAACGGCTGTACCA	1140
Db	4215	ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTAAACGGCTGTACCA	4274
Qy	1141	ATTTCTTTTGTCTTTGGGTATATACATTAAACCTTAAATTTT	1181
Db	4275	ATTTCTTTTGTCTTTGGGTATATATTTAAACCTTAAACCA	4315

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	84.7%	Score 1000.2	DB 2	Length 9859
	Best Local Similarity	90.4%	Pred. No. 6.7e-294		
	Matches 1068	Conservative 0	Mismatches 113	Indels 0	Gaps 0
QY	1	ATGGGGCAATCTTTCCACCAGCAATCTTGGGATTTCTTTCCGACACCACTTGGAT	60		



QY 661 TCCAACTACTCAACCACTCTCTCTCTCAACTTGTCTGTTATGCTGATGTCTG 720  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1735 TCCAACTACTCAACCACTCTCTCTCTCAACTTGTCTGTTATGCTGATGTCTG 1794  
QY 721 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 781 CTGAGATATCAAGGTATGTTGCGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 840  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1855 CTGAGATATCAAGGTATGTTGCGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 1914  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 841 ACGGAGCAATGACAGAGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1915 ACGGAGCAATGACAGAGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1974  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 901 TGTCTGTACAAAACCTTGGATGGAACCTGACCTGTATTCCTCATCTCTCTCTCT 960  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1975 TGTCTGTACAAAACCTTGGATGGAACCTGACCTGTATTCCTCATCTCTCTCTCT 2034  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 961 TTGCGAAAATTTCTTATGAGATGAGGCTCTGAGCCGTTTCTCTCTCTCTCTCTCT 1020  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2035 TTGCGAAAATTTCTTATGAGATGAGGCTCTGAGCCGTTTCTCTCTCTCTCTCTCT 2094  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1021 CCATTGTCTCAGTGTCTCTGAGGCTTTCCGCCAGTGTCTCTCTCTCTCTCTCTCT 1080  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2095 CCATTGTCTCAGTGTCTCTGAGGCTTTCCGCCAGTGTCTCTCTCTCTCTCTCTCT 2154  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1081 ATGTTTACTGAGGAGGCAAGTCTGTACACCATCTTGAAGTCTCTCTCTCTCTCTCT 1140  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2155 ATGTTTACTGAGGAGGCAAGTCTGTACACCATCTTGAAGTCTCTCTCTCTCTCTCT 2214  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1141 ATTTCTTTTGTCTTGGGTATCATTTAAACCTTAAATAA 1181  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2215 ATTTCTTTTGTCTTGGGTATCATTTAAACCTTAAATAA 2255  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

RESULT 14  
US-08-715-808-13  
; Sequence 13, Application US/08715808  
; Patent No. 5981274  
; GENERAL INFORMATION:  
; APPLICANT: Tyrrell, D. Lorne J.  
; APPLICANT: Chaisomchit, Sumonta  
; TITLE OF INVENTION: Recombinant Hepatitis Virus Vectors  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/715,808  
; FILING DATE: 18-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: CHANG-02441  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:

QY 661 TCCAACTACTCAACCACTCTCTCTCTCAACTTGTCTGTTATGCTGATGTCTG 720  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1735 TCCAACTACTCAACCACTCTCTCTCTCAACTTGTCTGTTATGCTGATGTCTG 1794  
QY 721 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 781 CTGAGATATCAAGGTATGTTGCGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 840  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1855 CTGAGATATCAAGGTATGTTGCGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 1914  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 841 ACGGAGCAATGACAGAGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1915 ACGGAGCAATGACAGAGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1974  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 901 TGTCTGTACAAAACCTTGGATGGAACCTGACCTGTATTCCTCATCTCTCTCTCT 960  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1975 TGTCTGTACAAAACCTTGGATGGAACCTGACCTGTATTCCTCATCTCTCTCTCT 2034  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 961 TTGCGAAAATTTCTTATGAGATGAGGCTCTGAGCCGTTTCTCTCTCTCTCTCTCT 1020  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2035 TTGCGAAAATTTCTTATGAGATGAGGCTCTGAGCCGTTTCTCTCTCTCTCTCTCT 2094  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1021 CCATTGTCTCAGTGTCTCTGAGGCTTTCCGCCAGTGTCTCTCTCTCTCTCTCTCT 1080  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2095 CCATTGTCTCAGTGTCTCTGAGGCTTTCCGCCAGTGTCTCTCTCTCTCTCTCTCT 2154  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1081 ATGTTTACTGAGGAGGCAAGTCTGTACACCATCTTGAAGTCTCTCTCTCTCTCTCT 1140  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2155 ATGTTTACTGAGGAGGCAAGTCTGTACACCATCTTGAAGTCTCTCTCTCTCTCTCT 2214  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 1141 ATTTCTTTTGTCTTGGGTATCATTTAAACCTTAAATAA 1181  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 2215 ATTTCTTTTGTCTTGGGTATCATTTAAACCTTAAATAA 2255  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

Query Match 83.5%; Score 985.8; DB 2; Length 7463;  
Best Local Similarity 89.7%; Pred. No. 1.4e-289;  
Matches 1059; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 1 ATGAGGAGCAATCTTTCACCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 3675 ATGAGGAGCAATCTTTCACCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3734  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 61 CAGGCTTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 120  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 3735 CAGGCTTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 3794  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 121 TGGCCAGAGCCCAACCAAGTATGAGCTGAGAGCTTGGAGCTTGGAGCTTGGAGCTT 180  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 3795 TGGCCAGAGCCCAACCAAGTATGAGCTGAGAGCTTGGAGCTTGGAGCTTGGAGCTT 3854  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 181 GGAGGCTTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 240  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 3855 GGAGGCTTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 3914  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 241 CGGCTTCTCTTCACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 300  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 3915 CAGGCTTTCAGAGCAACCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 3974  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 301 TTGAGAAAACATCTCTTCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 360  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 3975 CTAAGAGACATCTCTTCAACCAATTCAGATTGGAGCTTCAATCCCAAGAGACAC 4034  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 361 CAAGATCCCAAGTATGAGCTGATATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4035 CAAGATCCCAAGTATGAGCTGATATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4094  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 421 AACCTGTTCGACATCT 480  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4095 AACCTGTTCGACATCT 4154  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 481 GCGGAGACATGAGAAATCATCATCAGATTCTCTAGAACCCCTCTCTCTCTCTCTCTCTCT 540  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4155 GTGAGAACATGAGAAATCATCATCAGATTCTCTAGAACCCCTCTCTCTCTCTCTCTCTCT 4214  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 541 GGGTTTCTTGTGACAAAGATTCCTCAATACCGAGCTGAGCTGAGTGTGAGT 600  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4215 GGGTTTCTTGTGACAAAGATTCCTCAATACCGAGCTGAGCTGAGTGTGAGT 4274  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 601 TCTCTCAATTTCTAGGGGGAATACCGTGTCTCTTGGCCAAATTTGGCAGTCTCCCAAC 660  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4275 TCTCTCAATTTCTAGGGGGAATACCGTGTCTCTTGGCCAAATTTGGCAGTCTCCCAAC 4334  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 661 TCCAACTACTCAACCACTCTCTCTCTCAACTTGTCTGTTATGCTGATGTCTG 720  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4335 TCCAACTACTCAACCACTCTCTCTCTCAACTTGTCTGTTATGCTGATGTCTG 4394  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 721 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4395 CGGCGTTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4454  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 781 CTGAGATATCAAGGTATGTTGCGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 840  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4455 CTGAGATATCAAGGTATGTTGCGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 4514  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 841 ACGGAGCAATGACAGAGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 4515 ACGGAGCAATGACAGAGCTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4574  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

QY 901 TGTCTGTACAAAACCTTGGATGGAACCTGACCTGTATTCCTCATCTCTCTCTCT 960  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-08-715-808-13

Db 4575 TGCGTCAAAACCTACGAGATGGAATGCACTGTATATCCATCCATCCTCTTG36CT 4634  
Qy 961 TTGGAAATTCCTATGGAGTGGGACCTCAGCCCGTTTCTCCGGCTCAGTTTACTAGT 1020  
Db 4635 TTCCCAAAATACCTATGGAGTGGGACCTCAGCTCCGTTTCTTGCTCAGTTTACTAGT 4694  
Qy 1021 CCATTGTTTCAGTGGTTCGTAGAGGCTTCCCACTGTTTGCTTCAGTTATATGAGATG 1080  
Db 4695 CCATTGTTTCAGTGGTTCGTAGAGGCTTCCCACTGTTTGCTTCAGTTATATGAGATG 4754  
Qy 1081 ATGTTGTAAGTGGGAGCCCAAGTCTGTACACCATCTTGAATCCCTTTTACCGCTGTACCA 1140  
Db 4755 ATGAGGATATGGGGCCCAAGCTGTACAGCATCTGTAGATTCCTTTATACCGCTGTACCA 4814  
Qy 1141 ATTTCTTTTGTCTTTGGGATATACATTTAAACCCATAATA 1181  
Db 4815 ATTTCTTTTGTCTCTGGGATATACATTTAAACCCATAACAA 4855

## RESULT 15

US-08-890-735C-1  
; Sequence 1, Application US/08890735C  
; Patent No. 6518014  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: HEPADNAVIRUS CORES  
; FILE REFERENCE: DC44A  
; CURRENT APPLICATION NUMBER: US/08/890,735C  
; CURRENT FILING DATE: 1997-07-11  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 3182  
; TYPE: DNA  
; ORGANISM: Hepatitis B Virus  
US-08-890-735C-1

Query Match 68.7%; Score 811.2; DB 4; Length 3182;  
Best Local Similarity 97.3%; Pred. No. 1.2e-236;  
Matches 825; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 334 AACTCCAACTTTCACCAAACTCTGCAAGATCCCAAGTGAAGGTCTGTATTTCCCT 393  
Db 1 AATTCACAACCTTTCACCAAACTCTGCAAGATCCCAAGTGAAGGTCTGTATTTCCCT 60  
Qy 394 GCTGGGCTCCATTCAGGAACAGTAACCCCTGTCGACTACTGCTCCCATATCG 453  
Db 61 GCTGGGCTCCATTCAGGAACAGTAACCCCTGTCGACTACTGCTCCCTATATCG 120  
Qy 454 TCAATCTTCTGAGAGATTGGGACCTGCGCGGAACATGAGAAATCATCATCAGATTC 513  
Db 121 TCAATCTTCTGAGAGATTGGGACCTGCGCGGAACATGAGAAATCATCATCAGATTC 180  
Qy 514 CTAGAGCCCTGCTGCTGTATACAGCGGGGTTTTTCTTGTGACAGAAATCTCAATA 573  
Db 181 CTAGAGCCCTGCTGCTGTATACAGCGGGGTTTTTCTTGTGACAGAAATCTCAATA 240  
Qy 574 CCGAGAGTCTAGACTGCTGCTGATCTCTCAATTTTCTAGGGGGAACATACCGTGT 633  
Db 241 CCGAGAGTCTAGACTGCTGCTGATCTCTCAATTTTCTAGGGGGAACATACCGTGT 300  
Qy 634 CTGGCCAAATTCGAGTCCCAACCTCAATCATCATCAACCTCTGTCTCAACT 693  
Db 301 CTGGCCAAATTCGAGTCCCAACCTCAATCATCATCAACCTCTGTCTCAACT 360  
Qy 694 TGTCTGCTTATCGCTGATGTGTCTGCGGCTTTTATCATCTTCTCTCATCTGCTG 753  
Db 361 TGTCTGCTTATCGCTGATGTGTCTGCGGCTTTTATCATCTTCTCTCATCTGCTG 420  
Qy 754 CTATGCGCATCTCTTGTGTTCTTCTGAGATATCAAGGATATGGCCGTTTGTCT 813  
Db 421 CTATGCGCATCTCTTGTGTTCTTCTGAGATATCAAGGATATGGCCGTTTGTCT 480

Qy 814 CTAATTCAGAGATCTTCAACACAGACGCGGACCATGACAGACCTGACT 873  
Db 481 CTAATTCAGAGATCTTCAACACAGACGCGGACCATGACAGACCTGACT 540  
Qy 874 CAAGAACTCTATGTATCTCTCTGCTGTACAAACCTTCGGATGGAATCTGAC 933  
Db 541 CAAGAACTCTATGTATCTCTCTGCTGTACAAACCTTCGGATGGAATCTGAC 600  
Qy 934 TGTATTCATCCATCATCTGAGGCTTTCGAAATTCCTATGGAGTGGGCTCAGCC 993  
Db 601 TGTATTCATCCATCATCTGAGGCTTTCGAAATTCCTATGGAGTGGGCTCAGCC 660  
Qy 994 CGTTTCCTGCTCAGTTTACTAGTCCATTTGTCAGTGTTCGTAGGGCTTTCCCC 1053  
Db 661 CGTTTCCTGCTCAGTTTACTAGTCCATTTGTCAGTGTTCGTAGGGCTTTCCCC 720  
Qy 1054 ACTGTTGGCTTCAGTTATATGAGATGTGTACTGGGGGCAAGTCTGTACACATC 1113  
Db 721 ACTGTTGGCTTCAGTTATATGAGATGTGTACTGGGGGCAAGTCTGTACACATC 780  
Qy 1114 TTGATCCCTTTTACCGCTGTATACCAATTTTCTTTGCTTGGGATATCATTTAAAC 1173  
Db 781 TTGATCCCTTTTACCGCTGTATACCAATTTTCTTTGCTTGGGATATCATTTAAAC 840  
Qy 1174 CTAATAA 1181  
Db 841 CTAACAA 848

Search completed: September 15, 2003, 02:59:48  
Job time : 76 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 15, 2003, 02:24:56 ; Search time 258 Seconds

(without alignments)  
11113.101 Million cell updates/sec

Title: US-09-821-877-1

Perfect score: 1181  
Sequence: 1 atggggcagaaattcttcac.....taccattaaaccttaataaa 1181Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1632420 seqs, 1213878141 residues

Total number of hits satisfying chosen parameters: 3264840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PC7_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PC7US_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1181	100.0	1181	10	US-09-821-877-1 Sequence 1, Appli
2	1128.2	95.5	8007	9	US-09-837-297-3 Sequence 3, Appli
3	1001.8	84.8	4084	11	US-09-781-891D-21 Sequence 21, Appli
4	1001.8	84.8	4496	11	US-09-781-891D-22 Sequence 22, Appli
5	962.2	81.5	1286	12	US-10-136-819-1 Sequence 1, Appli
6	962.2	81.5	1286	14	US-10-132-829-3 Sequence 3, Appli
7	813.2	68.9	8717	9	US-09-837-297-4 Sequence 4, Appli
8	811.2	68.7	3182	10	US-09-929-955-14 Sequence 14, Appli
9	811.2	68.7	3182	13	US-10-104-966-11 Sequence 14, Appli
10	810.8	66.7	846	10	US-09-247-890-11 Sequence 11, Appli
11	756.4	64.0	3221	11	US-09-848-616-133 Sequence 13, App
12	755.4	64.0	3215	14	US-10-209-264-1 Sequence 1, Appli
13	744.6	63.0	3215	14	US-10-142-358-1 Sequence 1, Appli
14	693	58.7	5618	14	US-09-821-877-6 Sequence 6, Appli
15	690	58.4	690	10	US-09-821-877-7 Sequence 7, Appli
16	675.6	57.2	690	10	US-09-821-877-7 Sequence 7, Appli

17	668.2	56.6	681	10	US-09-821-877-3 Sequence 3, Appli
18	659.8	55.9	5130	10	US-09-897-006-9 Sequence 9, Appli
19	659.8	55.9	5130	11	US-09-897-511A-9 Sequence 9, Appli
20	652.2	55.2	681	9	US-09-812-862-13 Sequence 13, Appli
21	643.4	54.5	1696	12	US-10-267-922-15 Sequence 15, Appli
22	633	53.6	1194	9	US-09-812-862-5 Sequence 5, Appli
23	604.2	51.2	681	15	US-10-169-668-1 Sequence 1, Appli
24	509.6	43.1	1056	9	US-09-812-862-3 Sequence 3, Appli
25	412	34.9	426	10	US-09-247-890-15 Sequence 15, Appli
26	411.6	34.9	426	11	US-09-781-891D-12 Sequence 12, Appli
27	411.6	34.9	426	11	US-09-781-891D-14 Sequence 14, Appli
28	411.2	34.8	550	14	US-10-260-451-17 Sequence 17, Appli
29	411.2	34.8	550	14	US-10-260-451-19 Sequence 19, Appli
30	410	34.7	426	11	US-09-781-891D-13 Sequence 13, Appli
31	406.8	34.4	426	11	US-09-781-891D-19 Sequence 19, Appli
32	405.2	34.3	426	11	US-09-781-891D-18 Sequence 18, Appli
33	394	33.4	426	11	US-09-781-891D-20 Sequence 20, Appli
34	393	33.3	426	11	US-09-781-891D-10 Sequence 10, Appli
35	390.8	33.1	426	11	US-09-781-891D-16 Sequence 16, Appli
36	390.8	33.1	426	11	US-09-781-891D-17 Sequence 17, Appli
37	387.6	32.8	426	11	US-09-781-891D-11 Sequence 11, Appli
38	384.4	32.5	426	11	US-09-781-891D-15 Sequence 15, Appli
39	382.8	32.4	426	11	US-09-781-891D-9 Sequence 9, Appli
40	359.4	30.4	522	15	US-10-169-668-3 Sequence 3, Appli
41	326.6	27.7	870	9	US-09-812-862-7 Sequence 7, Appli
42	320.2	27.1	7991	9	US-09-837-297-5 Sequence 5, Appli
43	287.6	24.4	1041	9	US-09-812-862-1 Sequence 1, Appli
44	195	16.5	235	9	US-09-837-297-1 Sequence 1, Appli
45	171	14.5	182	10	US-09-821-877-4 Sequence 4, Appli

## ALIGNMENTS

```
RESULT 1
US-09-821-877-1
; Sequence 1, Application US/09821877
; Patent No. US20020177124A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Coleman, Paul F.
; APPLICANT: Mushawar, Isa K.
; TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
; FILE REFERENCE: 6794.US.O1
; CURRENT APPLICATION NUMBER: US/09/821.877
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1181
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
US-09-821-877-1
```

Query Match	Best Local Similarity	Score	DB 10:	Length	1181:
100.0%	100.0%	Pred. No. 0;	Matches 1181; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGGGGCAGAACTCTTCCACGACGAACTCTGGGATCTTCCCGACACGATTGGAT	60		
DB	1	ATGGGGCAGAACTCTTCCACGACGAACTCTGGGATCTTCCCGACACGATTGGAT	60		
QY	61	CGAGCTTCAGACCAACCAACCAATCCAGATTGGACTTCAATCCCAACGAGACACC	120		
DB	61	CGAGCTTCAGACCAACCAACCAATCCAGATTGGACTTCAATCCCAACGAGACACC	120		
QY	121	TGGCCAGCGCCACCAAGTGGAGCTGGACATTTGGACTGGGGTTTACCCACCGCAC	180		
DB	121	TGGCCAGCGCCACCAAGTGGAGCTGGACATTTGGACTGGGGTTTACCCACCGCAC	180		
QY	181	GGAGGCTTTGGGGTGGAGCCCTCAGGCTCAGGCGCATTAACAACTTGCAGCAAT	240		
DB	181	GGAGGCTTTGGGGTGGAGCCCTCAGGCTCAGGCGCATTAACAACTTGCAGCAAT	240		

Db 181 GGAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGGCAATACAAACCTTGGCCAGCAAAAT 240  
QY 241 CGGCTCTCTGCTTCCACCAATCGCAGTTCAGGAAGGAGGCTACCCGCTGTCTCCACCT 300  
Db 241 CGGCTCTCTGCTTCCACCAATCGCAGTTCAGGAAGGAGGCTACCCGCTGTCTCCACCT 300  
QY 301 TTGAGAAACAATCCTCAAGCCATGCACTGAGAACTCCAACTTTCACCAAACTCTG 360  
Db 301 TTGAGAAACAATCCTCAAGCCATGCACTGAGAACTCCAACTTTCACCAAACTCTG 360  
QY 361 CAAGATCCCAAGTGAAGGTCTGTATTTCCCTGCTGCTGCTCACTTCAAGAACTGA 420  
Db 361 CAAGATCCCAAGTGAAGGTCTGTATTTCCCTGCTGCTGCTCACTTCAAGAACTGA 420  
QY 421 AACCTGTTCGCACTGCTGTCTCCATATCCGATCTTCTGAGAGATTGGGGACCT 480  
Db 421 AACCTGTTCGCACTGCTGTCTCCATATCCGATCTTCTGAGAGATTGGGGACCT 480  
QY 481 GCGCGAACAATGAGAAATCAATCAGATTCCTTGAAGCCCTGCTGTGTTACAGCG 540  
Db 481 GCGCGAACAATGAGAAATCAATCAGATTCCTTGAAGCCCTGCTGTGTTACAGCG 540  
QY 541 GGGTTTTTCTTTGACAAAGATCTCAAAATCCGAGAGTCAAGTCTGCTGCTGAGT 600  
Db 541 GGGTTTTTCTTTGACAAAGATCTCAAAATCCGAGAGTCAAGTCTGCTGCTGAGT 600  
QY 601 TCTCTCAATTTCTAGGGGGAACATACGCTGTCTCAAACTTGTCTGAGTATCGCTGAGTGTCTG 720  
Db 601 TCTCTCAATTTCTAGGGGGAACATACGCTGTCTCAAACTTGTCTGAGTATCGCTGAGTGTCTG 720  
QY 661 TCCATACCTACCAACCTCTGTCTCTCAAACTTGTCTGAGTATCGCTGAGTGTCTG 720  
Db 661 TCCATACCTACCAACCTCTGTCTCTCAAACTTGTCTGAGTATCGCTGAGTGTCTG 720  
QY 721 CGGCGTTTATCATCTTCTCTCTTATCTGCTGCTGCTGCTCACTTCTTGTGTTCTT 780  
Db 721 CGGCGTTTATCATCTTCTCTCTTATCTGCTGCTGCTGCTCACTTCTTGTGTTCTT 780  
QY 781 CTGACTATCAAGGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAAGC 840  
Db 781 CTGACTATCAAGGTATGTTGCCGTTTGTCTCTAATTCAGAGATCTTCAACCAAGC 840  
QY 841 ACGGACCAATGACAGGCTGACAGCTCTGCTCAAGAACTTCTATGTAATCCCTCTGT 900  
Db 841 ACGGACCAATGACAGGCTGACAGCTCTGCTCAAGAACTTCTATGTAATCCCTCTGT 900  
QY 901 TGTGTACAAACCTTGTGATGAACTGCACTGTATTCCTATCCATCATCTGAGCT 960  
Db 901 TGTGTACAAACCTTGTGATGAACTGCACTGTATTCCTATCCATCATCTGAGCT 960  
QY 961 TTGGGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGAGCTCAGTTATCTAGTG 1020  
Db 961 TTGGGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGAGCTCAGTTATCTAGTG 1020  
QY 1021 CCAATTTGATGAGTGTCTGATGAGGCTTCCCACTGTTTGGCTTCAAGTTATGAGTG 1080  
Db 1021 CCAATTTGATGAGTGTCTGATGAGGCTTCCCACTGTTTGGCTTCAAGTTATGAGTG 1080  
QY 1081 ATGTTTACTGAGGGGCAAGTCTGTACCAATCTTGAAGTCCCTTTTACCGCTGTACCA 1140  
Db 1081 ATGTTTACTGAGGGGCAAGTCTGTACCAATCTTGAAGTCCCTTTTACCGCTGTACCA 1140  
QY 1141 ATTTTCTTTTGTCTTTGGGTATACATTAAACCTTAATAA 1181  
Db 1141 ATTTTCTTTTGTCTTTGGGTATACATTAAACCTTAATAA 1181

## RESULT 2

US-09-837-297-3  
; Sequence 3, Application US/09837297  
; Patent No. US20010049143A1  
; GENERAL INFORMATION:  
; APPLICANT: RYU, WANG SHICK

; TITLE OF INVENTION: Hepatitis B virus vectors for gene therapy  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/837,297  
; CURRENT FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: KR2000-21070  
; PRIOR FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Kopatentlin 1.71  
; SEQ ID NO 3  
; LENGTH: 8007  
; TYPE: DNA  
; ORGANISM: HBV  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)-(8007)  
; OTHER INFORMATION: Prototype vector of HBV  
US-09-837-297-3

Query Match 95.5%; Score 1128.2; DB 9; Length 8007;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1148; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGGGCGAAGATCTTTTCCACAGCAATCTCTGAGATTCTTCCGACCACTGAGT 60  
Db 1031 ATGGGCGAAGATCTTTTCCACAGCAATCTCTGAGATTCTTCCGACCACTGAGT 1090  
QY 61 CCAAGCTTCAAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGAGACCC 120  
Db 1091 CCAAGCTTCAAGCAACCAACCAATCCAGATTGGGACTTCAATCCCAAGAGACCC 1150  
QY 121 TGGCAGACCGCAACAGGTAGAGCTGAGCAATTCGAGTGGGGTTCAACCCGAC 180  
Db 1151 TGGCAGACCGCAACAGGTAGAGCTGAGCAATTCGAGTGGGGTTCAACCCGAC 1210  
QY 181 GAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGCGATACAAACCTTGGCAAGAAAT 240  
Db 1211 GAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGCGATACAAACCTTGGCAAGAAAT 1270  
QY 241 CGGCTCTCTGCTTCCACCAATCGCAGTTCAGAAAGGAGGCTACCCGCTGTCTCACT 300  
Db 1271 CGGCTCTCTGCTTCCACCAATCGCAGTTCAGAAAGGAGGCTACCCGCTGTCTCACT 1330  
QY 301 TTGAGAAACAATCCTCAAGCCATGCACTGAGAACTTCCAACTTTCACCAAACTCTG 360  
Db 1331 TTGAGAAACAATCCTCAAGCCATGCACTGAGAACTTCCAACTTTCACCAAACTCTG 1390  
QY 361 CAAGATCCCAAGTGAAGGTCTGTATTTCCCTGCTGCTGCTCACTTCAAGAACTGA 420  
Db 1391 CAAGATCCCAAGTGAAGGTCTGTATTTCCCTGCTGCTGCTCACTTCAAGAACTGA 1450  
QY 421 AACCTGTTCGCACTGCTGTCTCCATATCCGATCTTCTGAGAGATTGGGGACCT 480  
Db 1451 AACCTGTTCGCACTGCTGTCTCCATATCCGATCTTCTGAGAGATTGGGGACCT 1510  
QY 481 GCGCGAACAATGAGAAATCAATCAGATTCCTTGAAGCCCTGCTGTGTTACAGCG 540  
Db 1511 GCGCGAACAATGAGAAATCAATCAGATTCCTTGAAGCCCTGCTGTGTTACAGCG 1570  
QY 541 GGGTTTTTCTTTGACAAAGATCTCAAAATCCGAGAGTCTTGAAGTCTGCTGAGT 600  
Db 1571 GGGTTTTTCTTTGACAAAGATCTCAAAATCCGAGAGTCTTGAAGTCTGCTGAGT 1630  
QY 601 TCTCTCAATTTCTAGGGGGAACATACGCTGTCTGAGCAATTCCTGAGAGATTGGGGACCT 660  
Db 1631 TCTCTCAATTTCTAGGGGGAACATACGCTGTCTGAGCAATTCCTGAGAGATTGGGGACCT 1690  
QY 661 TCCATACCTACCAACCTCTGTCTCAAACTTGTCTGAGTATCGCTGAGTGTCTG 720  
Db 1691 TCCATACCTACCAACCTCTGTCTCAAACTTGTCTGAGTATCGCTGAGTGTCTG 1750  
QY 721 CGGCGTTTATCATCTTCTCTCTTATCTGCTGCTGCTCACTTCTTGTGTTCTT 780  
Db 1751 CGGCGTTTATCATCTTCTCTCTTATCTGCTGCTGCTCACTTCTTGTGTTCTT 1810

QY 781 CTGACATCAAGTATGTTGCGGTTGTCTTAATTCAGAGATCTTCAACCAACGAC 840  
Db 1811 CTGACATCAAGTATGTTGCGGTTGTCTTAATTCAGAGATCTTCAACCAACGAC 1870  
QY 841 ACGGACCATGACAGCCTGACACGACTCTGCTCAAGAACCTCTATGATCCCTCCGT 900  
Db 1871 ACGGACCATGACAGCCTGACACGACTCTGCTCAAGAACCTCTATGATCCCTCCGT 1930  
QY 901 TGTCTACAAAACTTCCGATGGAATGCACTGATATCCATCCATCATCTGGGCT 960  
Db 1931 TGTCTACAAAACTTCCGATGGAATGCACTGATATCCATCCATCATCTGGGCT 1990  
QY 961 TTGGAAAAATTCCTATGAGAGTGGGCTCAGCCCGCTTCTCCTGGCTCAGTTTACTAGT 1020  
Db 1991 TTGGAAAAATTCCTATGAGAGTGGGCTCAGCCCGCTTCTCCTGGCTCAGTTTACTAGT 2050  
QY 1021 CCATTGTTCACTGTTGTTAGGCTTTTCCCACTGTTGGCTTTCACTATATGATG 1080  
Db 2051 CCATTGTTCACTGTTGTTAGGCTTTTCCCACTGTTGGCTTTCACTATATGATG 2110  
QY 1081 ATGTGATCTGGGGGCAAGTCTGTACACATCTGAGTCCCTTTTACCGCTTACCA 1140  
Db 2111 ATGTGATCTGGGGGCAAGTCTGTACACATCTGAGTCCCTTTTACCGCTTACCA 2170  
QY 1141 ATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 1181  
Db 2171 ATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 2211

RESULT 3  
US-09-781-891D-21  
; Sequence 21, Application US/09781891D  
; Publication No. US20030096222A1

; GENERAL INFORMATION:  
; APPLICANT: Delaney, William IV  
; APPLICANT: Locarnini, Stephen Alistair  
; APPLICANT: Chen, Robert Yung Ming  
; APPLICANT: Bartholomewsz, Angelina  
; APPLICANT: Isom, Harriet  
; TITLE OF INVENTION: An assay  
; FILE REFERENCE: 2378750/EJH  
; CURRENT APPLICATION NUMBER: US/09/781,891D  
; PRIORITY FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/179,948  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 21  
; LENGTH: 4084  
; TYPE: DNA  
; ORGANISM: HBV 1.28 genome  
; US-09-781-891D-21

Query Match 84.8%; Score 1001.8; DB 11; Length 4084;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 1069; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 ATGGGAGCAATCTTTCACGACGAAATCTCTGGGATCTTTCACGACGACGATGGAT 60  
Db 1436 ATGGGAGCAATCTTTCGTTCCCAACCTCTGGGATCTTTCACGACGATGGAT 1495  
QY 61 CCAGCTTCAGAGCAAAACCAACCAATCCAGTTGGGACTTAATCCCAACAGAGACAC 120  
Db 1496 CCGATTCGAGGCACTCAACCAATCCAGTTGGGACTTAATCCCAACAGAGACAC 1555  
QY 121 TGGCCAGACGCAACAGGTAGAGCTGAGCATTCGACCTGGGGTTCAACCCACGAC 180  
Db 1556 TGGCCAGACGCAACAGGTAGAGCTGAGCATTCGACCTGGGGTTCAACCCACGAC 1615  
QY 181 GAGGCGCTTTTGGGTGAGCGCTCAGGCTCAGGGCATTAACCAACCTTCCAGCAAT 240  
Db 1616 GAGGCGCTTTTGGGTGAGCGCTCAGGCTCAGGGCATTAATTAACCAACGATTCACCAAT 1675

QY 241 CCGCCTCTGCTTCACCAATGCGAGTCAGAGAGGAGCCTAACCCCGCTGTCCACCT 300  
Db 1676 CCGCCTCTGCTTCACCAATGCGAGTCAGAGAGGAGCCTAACCCCGCTGTCCACCT 1735  
QY 301 TTGAGAAACATCATCTCTCAAGCCATGAGTGAATCTCAACAATCTTCCAAACTCTG 360  
Db 1736 TTGAGAAACATCATCTCTCAAGCCATGAGTGAATCTCAACAATCTTCCAAACTCTG 1795  
QY 361 CAAGATCCAGAGTGAAGGTCTGTATTCCTGCTGAGGCTCCAGTTACAGAAACAGTA 420  
Db 1796 CAAGATCCAGAGTGAAGGTCTGTATTCCTGCTGAGGCTCCAGTTACAGAAACAGTA 1855  
QY 421 AACCTGTTCGACTACTGTCTCTCCATATGTCATCTTCTCGAGATTTGGGAGCCT 480  
Db 1856 AACCTGTTCGAAATTTGCTCTCAATCTGTCTAATCTCCGAGAGAGTGGGAGCCT 1915  
QY 481 GCGCGAAATGAGAGACATCATCAGGATTCCTAGAACCCCTGCTGTGTTACAGGCG 540  
Db 1916 GCGCGAAATGAGAGACATCATCAGGATTCCTAGAACCCCTGCTGTGTTACAGGCG 1975  
QY 541 GGGTTTTCTTTGACAAAGATCTCTCAATACCGAGAGTCTGACTCCGTGTGAGCT 600  
Db 1976 GGGTTTTCTTTGACAAAGATCTCTCAATACCGAGAGTCTGACTCCGTGTGAGCT 2035  
QY 601 TCTCTCAATTTCTTGGGGGAACTACCGTGTGTCTTGGCCAAATTTCCAGTCCCAAC 660  
Db 2036 TCTCTCAATTTCTTGGGGGAACTACCGTGTGTCTTGGCCAAATTTCCAGTCCCAAC 2095  
QY 661 TCCAAATCACTACCAACCTCTGCTCTCAATCTTGTCTGATGCTGTGATGCTGTG 720  
Db 2096 TCCAAATCACTACCAACCTCTGCTCTCAATCTTGTCTGATGCTGTGATGCTGTG 2155  
QY 721 CCGGCTTTTATCATTTCTCTCTTCACTCTGCTGTCTATGCTCTCATCTTCTGTGGTCT 780  
Db 2156 CCGGCTTTTATCATTTCTCTCTTCACTCTGCTGTCTATGCTCTCATCTTCTGTGGTCT 2215  
QY 781 CTGACATCAAGGATGTTGCGGCTTGTCTCTTAACTTCAGAGATCTTCAACCAACGAC 840  
Db 2216 CTGACATCAAGGATGTTGCGGCTTGTCTCTTAACTTCAGAGATCTTCAACCAACGAC 2275  
QY 841 ACGGACCATGACAGCCTGACACGACTCTGCTCAAGAACCTCTATGATCCCTCTGT 900  
Db 2276 ACGGACCATGACAGCCTGACACGACTCTGCTCAAGAACCTCTATGATCCCTCTGT 2335  
QY 901 TGTCTACAAAACTTCCGATGGAATGCACTGATATCCATCCATCATCTGGGCT 960  
Db 2336 TGTCTACAAAACTTCCGATGGAATGCACTGATATCCATCCATCATCTGGGCT 2395  
QY 961 TTGGAAAAATTCCTATGAGAGTGGGCTCAGCCCGCTTCTCCTGGCTCAGTTTACTAGT 1020  
Db 2396 TTGGAAAAATTCCTATGAGAGTGGGCTCAGCCCGCTTCTCCTGGCTCAGTTTACTAGT 2455  
QY 1021 CCATTGTTCACTGTTGTTAGGCTTTTCCCACTGTTGGCTTTCACTATATGATG 1080  
Db 2456 CCATTGTTCACTGTTGTTAGGCTTTTCCCACTGTTGGCTTTCACTATATGATG 2515  
QY 1081 ATGTGATCTGGGGGCAAGTCTGTACAGCATCTGAGATCCCTTAATACCGCTGTACCA 1140  
Db 2516 ATGTGATCTGGGGGCAAGTCTGTACAGCATCTGAGATCCCTTAATACCGCTGTACCA 2575  
QY 1141 ATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 1181  
Db 2576 ATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 2616

RESULT 4  
US-09-781-891D-22  
; Sequence 22, Application US/09781891D  
; Publication No. US20030096222A1  
; GENERAL INFORMATION:  
; APPLICANT: Delaney, William IV  
; APPLICANT: Locarnini, Stephen Alistair

APPLICANT: Chen, Robert Yung Ming  
APPLICANT: Bartholomew, Angeline  
APPLICANT: Isom, Harriet  
TITLE OF INVENTION: An assay  
FILE REFERENCE: 2378750/EJH  
CURRENT APPLICATION NUMBER: US/09/781,891D  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 60/179,948  
PRIOR FILING DATE: 2000-02-03  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 22  
LENGTH: 4496  
TYPE: DNA  
ORGANISM: HBV 1.5 genome  
US-09-781-891D-22

Query Match 84.8%; Score 1001.8; DB 11; Length 4496;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 1069; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

1 ATGGGGCAGATCTTTCCACGACATCTCTGGGATTTCTTCCGACACAGTTGGAT 60  
1848 ATGGGACGAAATCTTTCTGTTCCACCTCTGGGATTTCTTCCGATCATCAGTTGGAC 1907  
61 CCAGCCTTCAGAGCAACACCAACATCCAGATTGGACTTCAATCCCAACAGACACC 120  
1908 CTGCAATTCGAGCACTCAACCAATCCAGATTGGACTTCAACCCCATCAAGACAC 1967  
121 TGCCCAAGCCCAACAGATGAGCTGAGCATTTGGAGTTCCGGGTTCCACCCGAC 180  
1968 TGCCCAAGCCCAACAGATGAGCTGAGCATTTGGAGTTCCGGGTTCCACCCGAC 2027  
181 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATTAACAACTTGGCCAGCAAT 240  
2028 GGGGATTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATTAACAACTTGGCCAGCAAT 2087  
241 CCGGCTCTCTGCTTCCACCAATCGCAGTCAAGAGGAGCAGCTTACCCGCTGTCTCACT 300  
2088 CTTCTCTCTGCTTCCACCAATCGCAGTCAAGAGGAGCAGCTTACCCGCTGTCTCACT 2147  
301 TTGAGAAACATCTACTCTCAAGCCATGAGTGAACCTCCCAACTTTCCCAAACTCTG 360  
2148 CTAGAGACAGTCACTCTCAGGCTCAGTGAATTCCTGCTTCCCAAGCTCTG 2207  
361 CAAGATCCAGAGTGAAGTCTGTATTTCTCTGCTGCTGCTCAGTTCAAGAACAGTA 420  
2208 CAGATCCAGAGTGAAGTCTGTATTTCTCTGCTGCTGCTCAGTTCAAGAACAGTA 2267  
421 AACCTGTCTCGACTACTGTCTCTCCCATATCGTCAATCTTCTGAGGATTTGGGACCT 480  
2268 AACCTGTCTCGACTACTGTCTCTCCCATATCGTCAATCTTCTGAGGATTTGGGACCT 2327  
481 GCGGGAACATGAGAACATCAATCAGATTTCTTGAAGACCCCTGCTGTGTTCAGAGC 540  
2328 GTAGCAACATGAGAACATCAATCAGATTTCTTGAAGACCCCTGCTGTGTTCAGAGC 2387  
541 GGGTTTTTCTGTGACAAAGATCTCTCAATACCGAGAGTCTAGACTGTGTGACT 600  
2388 GGGTTTTTCTGTGACAAAGATCTCTCAATACCGAGAGTCTAGACTGTGTGACT 2447  
601 TCTCTCAATTTCTAGGGGGAACCTACCTGTGTCTTGGCCAAATTTGCGATCCCAAC 660  
2448 TCTCTCAATTTCTAGGGGGAACCTACCTGTGTCTTGGCCAAATTTGCGATCCCAAC 2507  
661 TCCATTAATCAACAACTCTGTCTTCCAACTGTCTGTGTATTCGTGATGTGTCTG 720  
2508 TCCATTAATCAACAACTCTGTCTTCCAACTGTCTGTGTATTCGTGATGTGTCTG 2567  
721 CCGGCTTTTATCATCTTCTTCAATCAGTGTGCTGATAGCTCATCTTCTTGTGATCTT 780  
2568 CCGGCTTTTATCATCTTCTTCAATCAGTGTGCTGATAGCTCATCTTCTTGTGATCTT 2627

781 CTGAGCTATCAAGTATGTTGCCCGTTTGCTCTTAATTCAGAGATCTTCAACACAGC 840  
2628 CTGATTAATCAAGTATGTTGCCCGTTTGCTCTTAATTCAGAGATCTTCAACACAGC 2687  
841 ACGGACCATGACAGCTGACAGCTCTGCTCAAGAACTCTATGATCCCTCTGT 900  
2688 ACGGACCATGACAGCTGACAGCTCTGCTCAAGAACTCTATGATCCCTCTGT 2747  
901 TGGTGTACAAACCTTCGAGTGAACCTGACCTGTATTTCCATCCCATCTGCTG 960  
2748 TGGTGTACAAACCTTCGAGTGAACCTGACCTGTATTTCCATCCCATCTGCTG 2807  
961 TTGGAATAATCTTATGAGAGTGGGCTCAGCCCTTTCTCTGCTCAGTTACTAGT 1020  
2808 TTGGAATAATCTTATGAGAGTGGGCTCAGCCCTTTCTCTGCTCAGTTACTAGT 2867  
1021 CCATTTGTCAGTGTGCTGAGAGGCTTTCCCCCACTTTGGCTTCAATATAGATG 1080  
2868 CCATTTGTCAGTGTGCTGAGAGGCTTTCCCCCACTTTGGCTTCAATATAGATG 2927  
1081 ATGTTGTAATGGGGGCAAGTCTGTACACATCTTGAATCCCTTTTACCGCTGTAC 1140  
2928 ATGTTGTAATGGGGGCAAGTCTGTACACATCTTGAATCCCTTTTACCGCTGTAC 2987  
1141 ATTTTCTTTTGTCTTTGGGTATCATTTAACTTAATAA 1181  
2988 ATTTTCTTTTGTCTTTGGGTATCATTTAACTTAATAA 3028

## RESULT 5

US-10-136-819-1  
Sequence 1, Application US/10136819  
Publication No. US20030166593A1  
GENERAL INFORMATION:  
APPLICANT: Hoshijima, Masahiko  
TITLE OF INVENTION: No. US20030166593A1-viral vesicle vector for cardiac specific ger  
FILE REFERENCE: 6627-PAL198  
CURRENT APPLICATION NUMBER: US/10/136,819  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: 60/287,423  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1286  
TYPE: DNA  
ORGANISM: Hepatitis B virus  
US-10-136-819-1

Query Match 81.5%; Score 962.2; DB 12; Length 1286;  
Best Local Similarity 89.3%; Pred. No. 1,7e-302;  
Matches 1055; Conservative 0; Mismatches 108; Indels 18; Gaps 1;

1 ATGGGGCAGATCTTTTCCACGACATCTCTGGGATTTCTTCCGACACAGTTGGAT 60  
119 ATGGGGCAGATCTTTTCTGTTCCCAATCTCTGGGATTTCTTCCCATCAACAGTTGGAC 178  
61 CCAGCCTTCAGAGCAACACCAACATTCAGATTGGGACTTCAATCCCAACAGAGCACC 120  
179 CTTGCTGTGGAGCCCACTTAACAAATCCAGATTGGGACTTCAACCCCAACAGAGTCAA 238  
121 TGCCCAAGCCCAACAGATGAGCTGAGATTTGGAATTCGAGCTGAGGGTTCAACCCGAC 180  
239 TGCCCAAGCCCAACAGATGAGCTGAGATTTGGAATTCGAGCTGAGGGTTCAACCCGAC 298  
181 GAGAGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATTAACAACTTGGCCAGCAAT 240  
299 GGGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCAATTAACAACTTGGCCAGCAAT 358  
241 CCGGCTCTCTGCTTCCACCAATGCGAGTCAAGAGGAGGCTGACCCGCTGTCTCACT 300  
359 CTTCTCTCTGCTTCCACCAATGCGAGTCAAGAGGAGGCTGACCCCATCTCTCCACT 418

QY	301	TTGGAAACATCTAATCCTCAAGCCATGAGTGAACCTCAACAATCTTCCACCAACTCTG	360
Db	419	CTAGAGACATCTAATCCTCAAGCCATGAGTGAATTCACAACATTCACCAACTCTG	478
QY	361	CAAGATCCAGATGAGAGGTCTGTATTTCCCTGCTGAGTGCCTCCAGTTACAGAACAGTA	420
Db	479	CTAGATCCCAAGATGAGAGGCTTATATTTCTCTGCTGAGCTCCAGTTCCGGAACAGTA	538
QY	421	AACCTGTTCGACTACTGTCTCTCCCATATCTGTCAATCTTCTGAGAGATTGGGAGCCCT	480
Db	539	AACCTGTTCGACTACTGTCTCTCCCATATCTTCTGAGAGATTGGGAGCCCT	580
QY	481	GCAGGAAACATGAGAAACATCAATCAGAGATTCTCAGAGACCCTGCTGTATACAGGCG	540
Db	581	GCACCGAACAATGAGAAACACAAACATCAGAGATTCTCAGAGACCCTGCTGTATACAGGCG	640
QY	541	GAGTTTTTCTGTGTGACAAAGATCTCTCAATACCGCAGAGTCTAGACTGTGTGAGACT	600
Db	641	GAGTTTTTCTGTGTGACAAAGATCTCTCAATACCAAGAGTCTAGACTGTGTGAGACT	700
QY	601	TCTCATATTTCTAGGGGGAACCTACGCTGTCTGTGGCCAAATGTGCAATCCCAACC	660
Db	701	TCTCATATTTCTAGGGGGAACACCAAGTCTGTGGCCAAATGTGCAATCCCAACC	760
QY	661	TCCAACTACTACCAACCTCTGTCCTCAACTTGTCTGTGTTATCGCTGATGTGTG	720
Db	761	TCCAACTACTACCAACCTCTGTCCTCAATTTGTCTGTGCTATCGCTGATGTGTG	820
QY	721	CGGCTTTATCATCTTCTCTTCATCTGCTGTAGCTCTCATCTTCTGTGTGTTCTT	780
Db	821	CGGCTTTATCATATTCCTCTTCATCTGCTGTAGCTCTCATCTTCTGTGTGTTCTT	880
QY	781	CTGACATATCAAGGATGTGTCCGCTTGTCTCTCAATTCAGAGATCTTCAACACACAGC	840
Db	881	CTGACATACCAAGGATGTGTCCGCTTGTCTCTCAATTCAGAGAACATCAACACACAGC	940
QY	841	ACGGAGCCATGACAGAGCTGTGACGACTCTGCTCAAGAACCTCATGTATCCCTCTGT	900
Db	941	ACGGAGCCATGACAGAGCTGTGACGACTCTGCTCAAGAACCTCATGTATCCCTCTGT	1000
QY	901	TGCTGTACAAACCTTGCAGATGAACCTGACACTGTATTTCCATCCATCATCTGTGACT	960
Db	1001	TGCTGTACAAACCTTGCAGAGGAACTGACACTGTATTTCCATCCATCATCTGTGACT	1060
QY	961	TTGGAAATTCCTATGGGAGTGGGCTCAGCCGTTCTCTGTGCTAGTTACTAGT	1020
Db	1061	TTGGCAATTCCTATGGGAGTGGGCTCAGCCGTTCTCTGTGCTAGTTACTAGT	1120
QY	1021	CCATTTGTTCAGTGTGTCTGAGGGCTTTCGCCCACTGTTTGGCTTTCAGTTATATGATG	1080
Db	1121	CCATTTGTTCAGTGTGTCTGAGGGCTTTCGCCCACTGTTTGGCTTTCAGTTATATGATG	1180
QY	1081	ATGTGTATCTGGGGGCAAGTCTGTACACATCTTGAAGTCCCTTTTACCGCTGTACCA	1140
Db	1181	ATGTGTATCTGGGGGCAAGTCTGTACAACTTGAAGTCCCTTTTACCTATATACCA	1240
QY	1141	ATTTTCTTTGTCTTGGGTATACATTTAAACCCTAATAA	1181
Db	1241	ATTTTCTTTGTCTTGGGTATACATTTAAATGTGAATGAA	1281

Query Match	81.5%; Score 962.2; DB 14; Length 1286;	
Best Local Similarity	89.3%; Pred. No. 1.7e-302;	
Matches 1055; Conservative	0; Mismatches 108; Indels 18; Gaps 1	
Query	1 ATGAGGAGCAATCTTTCCACGACGAACTCTTGAGATTTCTTCCGACCCAGTTGGAT	60
Db	119 ATGAGGAGCAATCTTTCTGTTCCCAATCCTCGGATTTCTTCCGATCAGCAGTTGGAC	178
Qy	61 CCAGCCTTCAAGCAAAACCAACAATCCAGATTGGGACTTCAATCCCAAGAGACACC	120
Db	179 CCTCGTTTGGAGCCAACTCAAAACATCAATTGGGACTTCAACCCCAAGAGATCAA	238
Qy	121 TGGCCAGAGCCCAACAAGGTAGAGCTGGAGCATTCGGAATGGGGTTCAACCCACCGAC	180
Db	239 TGGCCAGAGCCCAAAATCAGGTAGAGCCGGAGCATTCGGGCCAGGGTTCAACCCACGAC	298
Qy	181 GGAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATTAACAACAACCTTGGCAGCAAT	240
Db	299 GGGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCT	358
Qy	241 CCGCTCTCTGCTTCCACCAATCGCAGTCAAGAGGACGCTTACCCCGCTGTCTCACT	300
Db	359 CCTCTCTCTGCTTCCACCAATCGCAGTCAAGAGGACGCTTACCCCGCTGTCTCACT	418
Qy	301 TTGAGAAACATCAATCTTCAAGCAATGAGTGGAACTCACAACCTTCCACCAACTCTG	360
Db	419 CTAAGAGACAGTCAATCTTCAAGCAATGAGTGGAACTCACAACCTTCCACCAACTCTG	478
Qy	361 CAGATCCCAAGTGAAGAGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAAGAAACAGTA	420
Db	479 CTAAGTCCCAAGTGAAGAGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAAGAAACAGTA	538
Qy	421 AACCTGTTCGCACTACTGTCTTCCATATGCTCAATCTTTCGAGGATTGGGACCT	480
Db	539 AACCTGTTCGCACTACTGTCTTCCATATGCTCAATCTTTCGAGGATTGGGACCT	580
Qy	481 GCGGGAACATGAGGAATCATCATCAGAAATTCCTTAGAACCCCTGCTGTATTAAGGG	540
Db	581 GCACCGAACAAGGAGAACCAACATCAGAAATTCCTTAGAACCCCTGCTGTATTAAGGG	640
Qy	541 GGGTTTTTCTTGTGAACAAGATCCTCAATTAACCGAGAGCTTGAAGTGGAGT	600
Db	641 GGGTTTTTCTTGTGAACAAGATCCTCAATTAACCGAGAGCTTGAAGTGGAGT	700
Qy	601 TCTCTCAATTTCTAGGGGAACTAACGCTGTCTTGGCCAAATTTGCAAGTCCCAACC	660
Db	701 TCTCTCAATTTCTAGGGGAGCAACCAAGTGTCTTGGCCAAATTTGCAAGTCCCAACC	760
Qy	661 TCAATCACTAACCACTCTCTGTCTTCAACTTGTCTTGGTTATTCGCTGAGTGTGTCTG	720
Db	761 TCAATCACTAACCACTCTCTGTCTTCAACTTGTCTTGGTTATTCGCTGAGTGTGTCTG	820
Qy	721 CGGGTTTTTATCATCTTCTGTCAATCTGAGCTATGCTCATCTTCTGTGGTTCT	780
Db	821 CGGGTTTTTATCATCTTCTGTCAATCTGAGCTATGCTCATCTTCTGTGGTTCT	880
Qy	781 CTGAGCTATCAAGGATGTGCTCCGTTGTCTCTTAATTCAGAGATCTTCAACCAACGAC	840
Db	881 CTGAGCTATCAAGGATGTGCTCCGTTGTCTCTTAATTCAGAGATCTTCAACCAACGAC	940
Qy	841 ACGGAGCAATCAGAGCTTGAAGCTCTGCTCAAGAACTCTATGATTCCTTCTGT	900

Db 941 ACCGGGCCATGCAAGACCTGACGATTCCTGCTCAAGAACTCTAATGTTCCCTCTTGT 1000  
QY 901 TGGTGTACAAACCTTGGATGTAAGTGAACCTGTATTCATCCCATGATCCTGGGCT 960  
Db 1001 TGGTGTACAAACCTTGGATGTAAGTGAACCTGTATTCATCCCATGATCCTGGGCT 1060  
QY 961 TTGGGAAATTTCCATATGGAGTGGGCTCAGCCGTTTCTCTGGCTCAGTTTATCTAGTG 1020  
Db 1061 TTGGCAGAAATTCCTATATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTATCTAGTG 1120  
QY 1021 CCATTGTTCAGTGTCTGTAGAGGCTTCCGCCCATGTTGGGCTTTGAGTTATATGATG 1080  
Db 1121 CCATTGTTCAGTGTCTGTAGAGGCTTCCGCCCATGTTGGGCTTTGAGTTATATGATG 1180  
QY 1081 ATGTTGTACTGGGGCCAAAGTCTGTACCACTTTGAGTCCCTTTTTCACCGCTGTACCA 1140  
Db 1181 ATGAGTATGGGGCCAAAGTCTGTACCACTTTGAGTCCCTTTTTCACCGCTGTATTCACCA 1240  
QY 1141 ATTTTCTTTTGTCTTTGGGTATATCTTTAAACCTTAATMAA 1181  
Db 1241 ATTTTCTTTTGTCTTTGGGTATATCTTTAAATGAAATTTGAA 1281

## RESULT 7

US-09-837-297-4  
; Sequence 4, Application US/09837297  
; Patent No. US20010049145A1  
; GENERAL INFORMATION:  
; APPLICANT: RYU, WANG SHICK  
; TITLE OF INVENTION: Hepatitis B virus vectors for gene therapy  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/837,297  
; CURRENT FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: KR2000-21070  
; PRIOR FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Koparentlin 1.71  
; SEQ ID NO 4  
; LENGTH: 8717  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: R711: pCMV-HBV/GFP Full Sequence  
US-09-837-297-4

Query Match 68.9%; Score 813.2; DB 9; Length 8717;  
Best Local Similarity 97.3%; Pred. No. 1,7e-253;  
Matches 827; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 332 GGAATTCACAACTTTTCACCAAACTCTGCAAGATCCCAAGTGAAGGTCTGTATTTC 391  
Db 2072 GGAATTCACAACTTTTCACCAAACTCTGCAAGATCCCAAGTGAAGGTCTGTATTTC 2131  
QY 392 CTGTGTGCTGCTCAGTTGAGAAAGTAACCTGTTCCGATCACTGCTCCCATAT 451  
Db 2132 CTGTGTGCTGCTCAGTTGAGAAAGTAACCTGTTCCGATCACTGCTCCCATAT 2191  
QY 452 CGTCAATCTTCTGAGAGATTGGGAGCCCTGCGGGAACATGAGAACATCAATCAGAT 511  
Db 2192 CGTCAATCTTCTGAGAGATTGGGAGCCCTGCGGGAACATGAGAACATCAATCAGAT 2251  
QY 512 TCCTAGAGACCCCTGCTGCTGTATACAGGCGGGGTTTCTTGTGACAAAGATCCTCAAA 571  
Db 2252 TCCTAGAGACCCCTTCTGCTGTATACAGGCGGGGTTTCTTGTGACAAAGATCCTCAAA 2311  
QY 572 TACCGAGAGTCTAGATCGTGTGGAATCTCTCAATTTTCTAGGGGGAACATACCGTGT 631  
Db 2312 TACCGAGAGTCTAGATCGTGTGGAATCTCTCAATTTTCTAGGGGGAACATACCGTGT 2371  
QY 632 GTCTTGGCCAAATTTGCAATGCCCAACTCAATCAATCAATCAATCAATCAATCAATCA 691  
Db 2372 GTCTTGGCCAAATTTGCAATGCCCAACTCAATCAATCAATCAATCAATCAATCAATCA 2431

QY 692 CTGTGCTGGTATTCGTGTGATGTGTGGGCGTTTATATCATCTTCTCTTCATCTGCG 751  
Db 2432 CTGTGCTGGTATTCGTGTGATGTGTGGGCGTTTATATCATCTTCTCTTCATCTGCG 2491  
QY 752 TGTATAGCCCACTCTTGTGTTGTTCTTCTGTGACATATCAAGTATGTTGCCGTTGTC 811  
Db 2492 TGTATAGCCCACTCTTGTGTTGTTCTTCTGTGACATATCAAGTATGTTGCCGTTGTC 2551  
QY 812 CTCTAATTCAGAGATCTTCAACACGAGACGGAACATGACAGCTGACACTCTG 871  
Db 2552 CTCTAATTCAGAGATCTTCAACACGAGACGGAACATGACAGCTGATGACTACTG 2611  
QY 872 CTCAAGAACTCTATGATATCCCTCTGTGCTGTACAAACCTTGGATGAACTGCA 931  
Db 2612 CTCAAGAACTCTATGATATCCCTCTGTGCTGTACAAACCTTGGAGGAAATTTG 2671  
QY 932 CCGTATTCACCATCATATCCTGGGCTTTGGGAAATTCCTATAGGAGTGGGCTCGAG 991  
Db 2672 CCGTATTCACCATCATATCCTGGGCTTTGGGAAATTCCTATAGGAGTGGGCTCGAG 2731  
QY 992 CCGGTTTCTCGGCTCAGTTTACTAGTGCATTTGTTCAGTGTTCGTAGGCTTTGCC 1051  
Db 2732 CCGGTTTCTCGGCTCAGTTTACTAGTGCATTTGTTCAGTGTTCGTAGGCTTTGCC 2791  
QY 1052 CCACTGTTGCTTTAGTATATGATATGATATGTTGACTGGGGCCAACTCTGTACCA 1111  
Db 2792 CCACTGTTGCTTTAGTATATGATATGATATGATATGATATGATATGATATGATATG 2851  
QY 1112 TCTTGAAGTCCCTTTTACCGGCTTACCAATTTCTTTGTTCTTGGGTATCATTTAA 1171  
Db 2852 TCTTGAAGTCCCTTTTACCGGCTTACCAATTTCTTTGTTCTTGGGTATCATTTAA 2911  
QY 1172 CCTTAATMAA 1181  
Db 2912 CCTTAATMAA 2921

## RESULT 8

US-09-929-955-14  
; Sequence 14, Application US/09929955  
; Patent No. US20020136740A1  
; GENERAL INFORMATION:  
; APPLICANT: Catharina Hulstgren  
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
; FILE REFERENCE: TRIPEP 23AUS2  
; CURRENT APPLICATION NUMBER: US/09/929,955  
; CURRENT FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: 09/705,547  
; PRIOR FILING DATE: 2000-11-03  
; PRIOR APPLICATION NUMBER: 60/229,175  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: 60/225,767  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 3182  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Hepatitis B virus sequence  
US-09-929-955-14

Query Match 68.7%; Score 811.2; DB 10; Length 3182;

Best Local Similarity 97.3%; Pred. No. 4.3e-253;  
Matches 825; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 334 AACTCACAAGCTTTCACCAAACTCTGCAAGATCCCAAGTGAAGGTGTGATTTCCCT 393  
Db 1 AATTCACAAGCTTTCACCAAACTCTGCAAGATCCCAAGTGAAGGTGTGATTTCCCT 60

```

QY 394 GCTGTGGCTCCAGTTCAGAAACAGTAAACCTGTGTCCAGTACTGTCTCCCATATCG 453
DB 61 GCTGTGGCTCCAGTTCAGAAACAGTAAACCTGTGTCTCCAGTACTGTCTCCCATATCG 120
QY 454 TCAATCTTCTCGAGATTGGGGACCTTGGCGGAACTGGAACATCATCAGATTTC 513
DB 121 TCAATCTTCTCGAGATTGGGGACCTTGGCGGAACTGGAACATCATCAGATTTC 180
QY 514 CTAGAACCCCTGCTCGTGTATTAAGGGGGGTTTTCTTGTGACAAAGATCTCAACATA 573
DB 181 CTAGAACCCCTGCTCGTGTATTAAGGGGGGTTTTCTTGTGACAAAGATCTCAACATA 240
QY 574 CCGCAGAGTCTAGACTCGTGTGAGACTTCTCAATTTTCTAGGGGAACTACCGTGTCT 633
DB 241 CCGCAGAGTCTAGACTCGTGTGAGACTTCTCAATTTTCTAGGGGAACTACCGTGTCT 300
QY 634 CTTGGCCAAATTTCCGAGTCCCAACCTCCATCACTACCAACCTCTCTCTCCACT 693
DB 301 CTTGGCCAAATTTCCGAGTCCCAACCTCCATCACTACCAACCTCTCTCTCCACT 360
QY 694 TGTCCGAGTATCGCTGATGTGTCTGGGGGTTTTATATCTCTCTCCATCCGCTG 753
DB 361 TGTCCGAGTATCGCTGATGTGTCTGGGGGTTTTATATCTCTCTCCATCCGCTG 420
QY 754 CTATGCTCATCTCTTGTGTGTTCTTCTGACATCAAGTATGTGGCCGTTGTCTCT 813
DB 421 CTATGCTCATCTCTTGTGTGTTCTTCTGACATCAAGTATGTGGCCGTTGTCTCT 480
QY 814 CTAAATTCAGATCTTCAACCAACGACGGAACCAATGACAGACTGACACTCTGCT 873
DB 481 CTAAATTCAGATCTTCAACCAACGACGGAACCAATGACAGACTGACACTCTGCT 540
QY 874 CAAGGAACCTATGATATCCCTGTTGCTGATCAAAACCTTGGATGGAACCTGACAC 933
DB 541 CAAGGAACCTATGATATCCCTGTTGCTGATCAAAACCTTGGATGGAACCTGACAC 600
QY 934 TGTATTTCCATCCATCATCATCTGAGGCTTTCGAAAAATTCCTATGGAGTGGCCCTGAC 993
DB 601 TGTATTTCCATCCATCATCATCTGAGGCTTTCGAAAAATTCCTATGGAGTGGCCCTGAC 660
QY 994 CGTTTCTCTGCTCAGTTTACTAGTGCATTTGTCAGTGTGCTGAGGGCTTTTCCCC 1053
DB 661 CGTTTCTCTGCTCAGTTTACTAGTGCATTTGTCAGTGTGCTGAGGGCTTTTCCCC 720
QY 1054 ACTGTTGGCTTTCAGTTATATGATGATGTGTGCTGGGGGCAAGCTGTACACCATC 1113
DB 721 ACTGTTGGCTTTCAGTTATATGATGATGTGTGCTGGGGGCAAGCTGTACACCATC 780
QY 1114 TTGAGTCCCTTTTACCGCTGTACCAATTTTCTTTGTCTTGGGTATACATTTAAAC 1173
DB 781 TTGAGTCCCTTTTACCGCTGTACCAATTTTCTTTGTCTTGGGTATACATTTAAAC 840
QY 1174 CTAAATPAA 1181
DB 841 CTAAACAA 848

```

```

RESULT 9
US-10-104-966-14
; Sequence 14, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matsi Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP. 23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104, 966
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705, 547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229, 175
; PRIOR FILING DATE: 2000-08-29

```

```

; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3182
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis B virus sequence
US-10-104-966-14

Query Match      68.7%; Score 811.2; DB 13; Length 3182;
Best Local Similarity 97.3%; Pred. No. 4,3e-253;
Matches 825; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 334 AACTCCACAACCTTCCACCAACCTGCAAGATCCCAAGTGAAGGTGTGATTTCCCT 393
DB 1 AATTCACAACCTTCCACCAACCTTGCAGAGATCCCAAGTGAAGGCTGTATTTCCCT 60
QY 394 GCTGTGGCTCCAGTTCAGAAACAGTAAACCTGTTCGAGTACTGTCTTCCCATATCG 453
DB 61 GCTGTGGCTCCAGTTCAGAAACAGTAAACCTGTTCGAGTACTGTCTTCCCATATCG 120
QY 454 TCAATCTTCTCGAGATTGGGGACCTTCGCGGAAACATGAGAACATCATCAGATTTC 513
DB 121 TCAATCTTCTCGAGATTGGGGACCTTCGCGTGAACATGAGAACATCATCAGATTTC 180
QY 514 CTAGAACCCCTGCTCGTGTATTAAGGGGGGTTTTCTTGTGACAAAGATCTCAACATA 573
DB 181 CTAGAACCCCTGCTCGTGTATTAAGGGGGGTTTTCTTGTGACAAAGATCTCAACATA 240
QY 574 CCGCAGAGTCTAGACTCGTGTGAGACTTCTCAATTTTCTAGGGGAACTACCGTGTCT 633
DB 241 CCGCAGAGTCTAGACTCGTGTGAGACTTCTCAATTTTCTAGGGGAACTACCGTGTCT 300
QY 634 CTTGGCCAAATTTCCGAGTCCCAACCTCCATCACTACCAACCTCTCTCTCCACT 693
DB 301 CTTGGCCAAATTTCCGAGTCCCAACCTCCATCACTACCAACCTCTCTCTCCACT 360
QY 694 TGTCCGAGTATCGCTGATGTGTCTGGGGGTTTTATATCTCTCTCCATCCGCTG 753
DB 361 TGTCCGAGTATCGCTGATGTGTCTGGGGGTTTTATATCTCTCTCCATCCGCTG 420
QY 754 CTATGCTCATCTCTTGTGTGTTCTTCTGACATCAAGTATGTGGCCGTTGTCTCT 813
DB 421 CTATGCTCATCTCTTGTGTGTTCTTCTGACATCAAGTATGTGGCCGTTGTCTCT 480
QY 814 CTAAATTCAGATCTTCAACCAACGACGGAACCAATGACAGACTGACACTCTGCT 873
DB 481 CTAAATTCAGATCTTCAACCAACGACGGAACCAATGACAGACTGACACTCTGCT 540
QY 874 CAAGGAACCTATGATATCCCTGTTGCTGATCAAAACCTTGGATGGAACCTGACAC 933
DB 541 CAAGGAACCTATGATATCCCTGTTGCTGATCAAAACCTTGGATGGAACCTGACAC 600
QY 934 TGTATTTCCATCCATCATCATCTGAGGCTTTCGAAAAATTCCTATGGAGTGGCCCTGAC 993
DB 601 TGTATTTCCATCCATCATCATCTGAGGCTTTCGAAAAATTCCTATGGAGTGGCCCTGAC 660
QY 994 CGTTTCTCTGCTCAGTTTACTAGTGCATTTGTCAGTGTGCTGAGGGCTTTTCCCC 1053
DB 661 CGTTTCTCTGCTCAGTTTACTAGTGCATTTGTCAGTGTGCTGAGGGCTTTTCCCC 720
QY 1054 ACTGTTGGCTTTCAGTTATATGATGATGTGTGCTGGGGGCAAGCTGTACACCATC 1113
DB 721 ACTGTTGGCTTTCAGTTATATGATGATGTGTGCTGGGGGCAAGCTGTACACCATC 780
QY 1114 TTGAGTCCCTTTTACCGCTGTACCAATTTTCTTTGTCTTGGGTATACATTTAAAC 1173
DB 781 TTGAGTCCCTTTTACCGCTGTACCAATTTTCTTTGTCTTGGGTATACATTTAAAC 840
QY 1174 CTAAATPAA 1181
DB 841 CTAAACAA 848

```



```
RESULT 10
US-09-247-890-11
; Sequence 11, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juhua
; APPLICANT: Baas, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247, 890
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Hepatitis B virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(846)
; OTHER INFORMATION: Pres2-S coding region of hepatitis B virus aYW
; OTHER INFORMATION: surface antigen (HBsAg)
US-09-247-890-11
```

```
Query Match      68.7%; Score 810.8; DB 10; Length 846;
Best Local Similarity 97.4%; Pred. No. 2.9e-253;
Matches 824; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```
QY      325 ATGCAGTGAAGTCCACAACTTTCACAACTCTGCAGATCCAGAGTGAAGTCTG 384
      1 ATGCAGTGAAGTCCACAACTTTCACAACTCTGCAGATCCAGAGTGAAGTCTG 60
QY      385 TATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
      61 TATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY      445 CCCATATCGTCATCTTCTGAGAGATGGGAGCCCTGCGGGAGGAACATGAGATCACA 504
      121 CCTTATCGTCATCTTCTGAGAGATGGGAGCCCTGCGGGAGGAACATGAGATCACA 180
QY      505 TCAGATTCCTAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
      181 TCAGATTCCTAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY      565 CTACACATACCGCAGAGTCTAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624
      241 CTACACATACCGCAGAGTCTAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY      625 ACCGTGCTCTTGGCCAAAATTCGAGTCCCAACCTCCAACTCATCCAACTCCTCT 684
      301 ACCGTGCTCTTGGCCAAAATTCGAGTCCCAACCTCCAACTCATCCAACTCCTCT 360
QY      685 CCTCCAACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
      361 CCTCCAACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY      745 ATCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
      421 ATCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY      805 GTTTGCTGCTATTCAGAGTCTTCAACGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 864
      481 GTTTGCTGCTATTCAGAGTCTTCAACGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
```

```
QY      865 ACTCTGCTCAAGAACCTTATGATCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 924
      541 ACTAGCTCAAGAACCTTATGATCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY      925 AACTGACCTGATATTCATCCATCCATCATCTGAGGCTTTGGGAAAATTCCTAT 984
      601 AATTGACCTGATATTCATCCATCCATCATCTGAGGCTTTGGGAAAATTCCTAT 660
QY      985 GCTCAGCCGCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
      661 GCTCAGCCGCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY      1045 CTTTCCCACTGTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGAT 1104
      721 CTTTCCCACTGTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGAT 780
QY      1105 TACACATCTTGAGTCCCTTTTACCGCTGTTACCAATTTCTTTGCTGCTG 1164
      781 TACACATCTTGAGTCCCTTTTACCGCTGTTACCAATTTCTTTGCTGCTGCTGCT 840
QY      1165 ATTAA 1170
      841 ATTAA 846
Db
```

```
RESULT 11
US-09-247-890-9
; Sequence 9, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juhua
; APPLICANT: Baas, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247, 890
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Hepatitis B virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(846)
; OTHER INFORMATION: Pres2-S coding region of hepatitis B virus aYr
; OTHER INFORMATION: surface antigen (HBsAg)
US-09-247-890-9
```

```
Query Match      64.0%; Score 756.4; DB 10; Length 846;
Best Local Similarity 93.4%; Pred. No. 1.5e-235;
Matches 790; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
```

```
QY      325 ATGCATGGAAGTCCCAACTTTCACAACTCTGCAAGATCCAGATGAGAGTCTG 384
      1 ATGCATGGAAGTCCCAACTTTCACAACTCTGCAAGATCCAGATGAGAGTCTG 60
QY      385 TATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
      61 TATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY      445 CCCATATCGTCATCTTCTCGAGAGATTTGGGAGCCCTGCGCGGAGCAATGAGAAATCACA 504
      121 CCCATATCGTCATCTTCTCGAGAGATTTGGGAGCCCTGCGCGGAGCAATGAGAAATCACA 180
Db
```





Lim, Gek Keow  
Zhao, Yi  
Chen, Wei Ning  
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND  
USSES THEREOF  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ladas & Parry  
STREET: 26 West 61 Street  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/209,264  
FILING DATE: 31-Jul-2002  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SG98/00046  
FILING DATE: 19-JAN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mass, Clifford J.  
REGISTRATION NUMBER: 30,086  
REFERENCE/DOCKET NUMBER: U-013109-7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 708-1800  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
US-10-209-264-1

Query Match 63.0%; Score 744.6; DB 14; Length 3215;  
Best Local Similarity 94.0%; Pred. No. 2.2e-231;  
Matches 796; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

QY 336 CTCACAACTTTCACCAAACTCTGCAAGATCCAGAGTGAAGTCTGATTTCCCTGC 395  
DB 1 CTCACAACTTTCACCAAACTCTGCAAGATCCAGAGTGAAGTCTGATTTCCCTGC 60  
QY 396 TGGTGGCTCCAGTTCAGGAACAGTAAACCTGTTCCGACTAATGCTCTCCCATATGCTC 455  
DB 61 TGGTGGCTCCAGTTCAGGAACAGTAAACCTGTTCCGACTAATGCTCTCCCATATGCTC 120  
QY 456 AATCTTCTCGAGATTTGGGAGCCCTGCGCGGAAACATGAGAAACATCAATCAGGATTCCT 515  
DB 121 AATCTTCTCGAGATTTGGGAGCCCTGCGCGGAAACATGAGAAACATCAATCAGGATTCCT 180  
QY 516 AGAACCCCTGCTGCTGTTAAGAGGCGGCTTTCTTGTTGCAAGAAATCTTCACATACC 575  
DB 181 AGAACCCCTGCTGCTGTTAAGAGGCGGCTTTCTTGTTGCAAGAAATCTTCACATACC 240  
QY 576 GCAGAGCTGAGACTCGGTGGTGAATTTCTCAATTTTCTAGGGGGAACATACCGTGTG-TC 634  
DB 241 GCAGAGCTGAGACTC-TGGTGGTGAATTTCTCAATTTTCTAGGGGGAACATACCGTGTTC 299  
QY 635 TTGGCCAAATTTGCGATGCCCAACCTCAATCACTGACCAACCTCTGTCTCCCACTT 694  
DB 300 CTGGCCAAATTTGCGATGCCCAACCTCAATCACTGACCAACCTCTGTCTCCCACTT 359  
QY 695 GTCTGTGTTATCGCTGATGTCTGCGGGGCTTTATCATCTTCTCTTCATCCTGCTGC 754  
DB 360 GTCTGTGTTATCGCTGATGTCTGCGGGGCTTTATCATCTTCTCTTCATCCTGCTGC 419  
QY 755 TATGCTCATCTCTTGTGTTGTTCTTCTGACATCAAGGATGTTGCCCGTTGTCTC 814

DB 420 TATGCTCATCTCTTGTGTTGTTCTTCTGACATCAAGGATGTTGCCCGTTGTCTC 479  
QY 815 TAAATCAGGATTTTCAACACGACGAGCATGACAGGCTGACAGCTCTGCTC 874  
DB 480 TAACTCCAGGACATCAACACGAGGAGCATGACAGCTGACAGCTCTGCTC 539  
QY 875 AAGGAACCTCTATGATATCCCTCTGTTGCTGTAACAAACCTTGATGAGAACTGACCT 934  
DB 540 AAGGAACCTCTATGATATCCCTCTGTTGCTGTAACAAACCTTGATGAGAACTGACCT 599  
QY 935 GTATTCATCCATCATCTGCGGCTTTGGGAAATTTCTATGAGAGTGGGCTCAGCC 994  
DB 600 GTATTCATCCATCATCTGCGGCTTTGGGAAATTTCTATGAGAGTGGGCTCAGTCC 659  
QY 995 GTTTCCTCCGCTCAGTTACTAGTGCATTTGTTGATGAGTGGTCTGAGGCTTTCCGCCA 1054  
DB 660 GTTTCCTCCGCTCAGTTACTAGTGCATTTGTTGATGAGTGGTCTGAGGCTTTCCGCCA 719  
QY 1055 CTGTTGGCTTTGAGTTATATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1114  
DB 720 CTGTTGGCTTTGAGTTATATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 779  
QY 1115 TGAGTCCCTTTTACCGCTGTACCAATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1174  
DB 780 TGAGTCCCTTTTACCGCTGTACCAATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 839  
QY 1175 TAAATAA 1181  
DB 840 TAAATAA 846

RESULT 14  
US-10-142-358-1  
Sequence 1, Application US/10142358  
Publication No. US20030083291A1  
GENERAL INFORMATION:  
APPLICANT: Michel, Marie-Louise  
Mancine, Maryline  
TITLE OF INVENTION: Nucleotide Vector, Composition  
Containing Such Vector, and Vaccine for Immunization  
Against Hepatitis  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/142,358  
FILING DATE: 10-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/799,569  
FILING DATE: 12-FEB-1997  
APPLICATION NUMBER: US 08/706,337  
FILING DATE: 30-AUG-1996  
APPLICATION NUMBER: US 08/633,821  
FILING DATE: 22-APR-1996  
APPLICATION NUMBER: FR 94/00483  
FILING DATE: 27-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03495, 0128-01000

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 5618 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-142-358-1

Query Match      58.7%; Score 693; DB 14; Length 5618;
Best Local Similarity 96.6%; Pred. No. 1.9e-214;
Matches 708; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 449 TATGCTCATCTTCTCGAGGATTGGGACCTCGCGGAACTGGAGAAATCATCATCAG 508
DB |||||
DB 898 TACCGGGCCCCCTCGAGGATTGGGGACCTCGCGTGAACATGAGAAATCATCATCAG 957
QY 509 GATTCCTAGGACCCCTGCTGCTGTTACAGGGGGGTTTTCTGTTGACAAAGATCCCA 568
DB |||||
DB 958 GATTCCTAGGACCCCTGCTGCTGTTACAGGGGGGTTTTCTGTTGACAAAGATCCCA 1017
QY 569 CAATACCCGAGAGTCTAGACTGCTGCTGCTCTCTCAATTTCTAGGGGAACTACCG 628
DB 1018 CAATACCCGAGAGTCTAGACTGCTGCTGCTCTCTCAATTTCTAGGGGAACTACCG 1077
QY 629 TGTGCTTGGGCAAAATTCGACAGTCCCACTCTCAATCATCTACCAACTCTGCTCTG 688
DB 1078 TGTGCTTGGGCAAAATTCGACAGTCCCACTCTCAATCATCTACCAACTCTGCTCTG 1137
QY 689 CAATTCCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
DB 1138 CAATTCCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197
QY 749 TGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
DB 1198 TGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1257
QY 809 GTCTCTAATTCGAGATCTTCAACCAACGAGACGAGACATGCAAGCTGACAGCTC 868
DB 1258 GTCTCTAATTCGAGATCTTCAACCAACGAGACGAGACATGCAAGCTGACAGCTC 1317
QY 869 CTGCTCAAGGAACCTCTATGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
DB 1318 CTGCTCAAGGAACCTCTATGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1377
QY 929 GCACCTGATTCGCAATCCCATCATCTGAGGCTTTGCGAAATTTCTATGGAGTGGGCT 988
DB 1378 GCACCTGATTCGCAATCCCATCATCTGAGGCTTTGCGAAATTTCTATGGAGTGGGCT 1437
QY 989 CAGCCGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
DB 1438 CAGCCGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1497
QY 1049 CCCCCACTGTTGGCTTTCAGTTATATGATGATGTTGATGCTGAGGGCCAAATCTGTACA 1108
DB 1498 CCCCCACTGTTGGCTTTCAGTTATATGATGATGTTGATGCTGAGGGCCAAATCTGTACA 1557
QY 1109 CCATCTGAGTCCCTTTTACCGCTGTTACCAATTTTCTTTGCTGCTTGGGATACATTT 1168
DB 1558 GCATTTTGAATCCCTTTTACCGCTGTTACCAATTTTCTTTGCTTGGGATACATTT 1617
QY 1169 AAACCTATATAA 1181
DB 1618 AAACCTATATAA 1630
```

RESULT 15  
US-09-821-877-6  
; Sequence 6, Application US/09821877

```
; Patent No. US20020177124A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Coleman, Paul F.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
; TITLE OF INVENTION: And Methods Of Detection Thereof
; FILE REFERENCE: 6794.US.01
; CURRENT APPLICATION NUMBER: US/09/821,877
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Hepatitis B Virus
; OTHER INFORMATION: Mutant Hepatitis B Virus Strain
US-09-821-877-6

Query Match      58.4%; Score 690; DB 10; Length 690;
Best Local Similarity 100.0%; Pred. No. 5.9e-214;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 GCGCGGAACATGAGAACATCATCAGATTCCTAGGACCCCTGCTGTTACAGGCG 540
DB 1 GCGCGGAACATGAGAACATCATCAGATTCCTAGGACCCCTGCTGTTACAGGCG 60
QY 541 GGGTTTTTCTGTTGACAAAGATCTCTCAATACCGCAGATCTAGACTCTGTTGAGCT 600
DB 61 GGGTTTTTCTGTTGACAAAGATCTCTCAATACCGCAGATCTAGACTCTGTTGAGCT 120
QY 601 TCTCTCAATTTCTAGGGGGAACTACCGTGTGCTGCGCAAAATTCGACGCCCCAAC 660
DB 121 TCTCTCAATTTCTAGGGGGAACTACCGTGTGCTGCGCAAAATTCGACGCCCCAAC 180
QY 661 TCCATCACTACCAACCTCTGCTCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 181 TCCATCACTACCAACCTCTGCTCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 721 CGGCGTTTATCATCTTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 241 CGGCGTTTATCATCTTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 781 CTGACCTATCAAGGATGTTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 301 CTGACCTATCAAGGATGTTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 841 ACGGACCATGACAGAGCTGACAGACTCTGCTCAAGGAACCTCTATGATCCCTCTGT 900
DB 361 ACGGACCATGACAGAGCTGACAGACTCTGCTCAAGGAACCTCTATGATCCCTCTGT 420
QY 901 TGTGTTACAAAACCTTGGATGAGAACTGACCTGATATTCATCCATCATCTGAGCT 960
DB 421 TGTGTTACAAAACCTTGGATGAGAACTGACCTGATATTCATCCATCATCTGAGCT 960
QY 961 TTGGAATAATTCCTATGAGAGTGGGCTCAGCCGCTTCTCTGCTGCTGCTGCTGCTG 1020
DB 481 TTGGAATAATTCCTATGAGAGTGGGCTCAGCCGCTTCTCTGCTGCTGCTGCTGCTG 540
QY 1021 CCATTTGTCAGTGTGCTGAGGGCTTCCCACTGTTTGGCTTCAAGTATATGAGATG 1080
DB 541 CCATTTGTCAGTGTGCTGAGGGCTTCCCACTGTTTGGCTTCAAGTATATGAGATG 600
QY 1081 ATGTTGATCGGGGCAAGTCTGTACACCAATCTTGAAGTCCCTTTTACCGGTATACA 1140
DB 601 ATGTTGATCGGGGCAAGTCTGTACACCAATCTTGAAGTCCCTTTTACCGGTATACA 660
QY 1141 ATTTCTTTTGTCTTGGGTATACATTTAA 1170
DB 661 ATTTCTTTTGTCTTGGGTATACATTTAA 690
```

Tue Sep 16 09:06:12 2003

us-09-821-877-1.rmpb

Page 12

Search completed: September 15, 2003, 03:56:35  
Job time : 263 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2003, 01:23:11; Search time 1985 Seconds  
(without alignments)  
14460.245 Million cell updates/sec

Title: US-09-821-877-1

Perfect score: 1181  
Sequence: 1 atggggcagaatcttcac.....tacatttaaccctaataa 1181

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estmu:\*  
4: em\_estnu:\*  
5: em\_estrov:\*  
6: em\_estropl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_ges\_hum:\*  
18: em\_ges\_hiv:\*  
19: em\_ges\_pin:\*  
20: em\_ges\_vrt:\*  
21: em\_ges\_fun:\*  
22: em\_ges\_mam:\*  
23: em\_ges\_mus:\*  
24: em\_ges\_pro:\*  
25: em\_ges\_rtd:\*  
26: em\_ges\_phg:\*  
27: em\_ges\_vrl:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	45.8	3.9	997	29	CNS005TE	AL060767 Drosophila
2	45.6	3.9	636	28	AZ038456	AZ038456 RPT-23-2
3	45.6	3.9	885	13	BX425603	BX425603 BX425603
4	43.2	3.7	779	29	BZ696852	BZ696852 SP_Ba008

C 5	42.6	3.6	602	28	AZ372526	AZ372526 1M0124L21
6	42.4	3.6	1201	13	BX381961	BX381961 BX381961
7	41.8	3.5	1101	29	CNS016LW	AL106910 Drosophila
8	41.6	3.5	308	28	AZ879026	AZ879026 RPT-23-1
9	41.5	3.5	377	28	AZ413108	AZ413108 1M0197017
C 10	41.3	3.5	745	28	BH034701	BH034701 RPT-24-2
C 11	40.8	3.4	697	28	BH063007	BH063007 RPT-24-3
C 12	40.6	3.4	629	9	AV647972	AV647972 AV647972
13	40.3	3.4	512	28	AZ938251	AZ938251 2M0196F16
14	40.3	3.4	520	28	BZ128636	BZ128636 CH230-440
C 15	40.3	3.4	1101	29	CNS016B8	AL106526 Drosophila
C 16	39.8	3.4	1254	29	AG072222	AG072222 Pan trogl
C 17	39.6	3.4	450	28	AZ097652	AZ097652 RPT-23-1
C 18	39.4	3.3	548	28	BH107673	BH107673 RPT-23-3
19	39.4	3.3	644	10	BB641359	BB641359 BB641359
C 20	39.4	3.3	681	29	CNS02FE9	AL194922 Tetraodon
C 21	39.4	3.3	1003	29	CNS01TT3	AL166844 Tetraodon
22	39.4	3.3	1201	13	BX337072	BX337072 BX337072
23	39.4	3.3	1653	11	AK043679	AK043679 Mus muscu
24	39.2	3.3	292	28	AZ637334	AZ637334 1M0496108
C 25	39.2	3.3	987	29	CNS00418	AL066537 Drosophila
C 26	39.2	3.3	1101	29	CNS0181N	AL108773 Drosophila
C 27	39.2	3.3	1201	13	BX376097	BX376097 BX376097
C 28	39.2	3.3	239	10	BF78098	BF78098 NXSL-077
C 29	39.2	3.3	555	28	AZ882512	AZ882512 RPT-23-2
C 30	39.2	3.3	621	12	BM492837	BM492837 NXRV-031
C 31	39.2	3.3	705	12	BM902701	BM902701 NXRV-057
C 32	39.2	3.3	972	29	CNS02BJT	AL189928 Tetraodon
C 33	39.2	3.3	985	9	AL575191	AL575191 AL575191
C 34	39.2	3.3	1036	12	BM011009	BM011009 603634682
35	39.2	3.3	1212	29	BZ560978	BZ560978 pac82-164
36	38.8	3.3	885	13	BX425603	BX425603 BX425603
37	38.6	3.3	597	28	AZ067459	AZ067459 RPT-23-4
C 38	38.6	3.3	611	14	CB580512	CB580512 AMGNNUC-C
C 39	38.4	3.3	668	28	AZ312109	AZ312109 1M0027K13
C 40	38.4	3.3	924	10	BF241268	BF241268 601878783
C 41	38.4	3.3	874	13	BX344914	BX344914 BX344914
42	38.4	3.3	1101	29	CNS00330	AL063790 Drosophila
43	38.2	3.2	214	9	AI053945	AI053945 q171h02.x
C 44	38.2	3.2	463	12	B1160947	B1160947 602865330
45	38.2	3.2	492	28	BZ180127	BZ180127 CH230-349

## ALIGNMENTS

RESULT 1  
CNS005TE 997 bp DNA linear GSS 03-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
DEFINITION BACR12K22 of RPT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL060767.1 GI:4943573  
VERSION AL060767  
KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 997)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oosagawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of





Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weis, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

## JOURNAL

Unpublished

Contact: Robert B. Weis

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., STC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0124 row: L column: 21

Seq primer: CGTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 602.

## FEATURES

source

Location/Qualifiers

1. .602

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UGC1M0124L21"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone\_lib="Mouse 10kb plasmid UGC1M library"

/note="Vector: PWD42nv, Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD42 (gii4732114[gb|AF129072.1]), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

## BASE COUNT

320 a 48 c 197 g 37 t

## ORIGIN

Query Match 3.6%; Score 42.6; DB 28; Length 602;  
Best Local Similarity 49.3%; Pred. No. 4.9;  
Matches 111; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY TCTCAATTTTCAGGGGAACACGCTGTGCTTGGCCAAATTCGAGTCCCAACCTC 662  
Db TCTCTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 167  
QY 663 CAATCATCAACCACTCTCTCTCAACTGTGCTGTTATGCGTGAATGTCTGCG 722  
Db 166 TTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 107  
QY 723 GCGTTTATCATCTTCTTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782  
Db 106 TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 47  
QY 783 GGAATATCAAGATGTGCGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827  
Db 46 CTCTCTCTCTTCTGCTTCTGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTG 2

RESULT 6

EX381961 1201 bp mRNA linear EST 08-MAY-2003  
LOCUS EX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS0D1072YF05 3-PRIME, mRNA sequence.

ACCESSION EX381961 GI:30453007

VERSION EX381961.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 1201)

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/InvitrogenCorporation 1600

Faraday Avenue Genoscope sequence ID : CS0D1072CC03NP1.

## FEATURES

source

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1072YF05"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

## BASE COUNT

95 a 191 c 115 g 55 t 745 others

## ORIGIN

Query Match 3.6%; Score 42.4; DB 13; Length 1201;  
Best Local Similarity 2.6%; Pred. No. 7.4;  
Matches 10; Conservative 111; Mismatches 264; Indels 0; Gaps 0;

QY 768 CTGTGGTCTTCTTCTGACATCAAGATGTGTCCTGCTTCTCTTAATTCAGATC 827  
Db 410 MNKK 469  
QY 828 TTGAACCAACGACGAGCAATGACGAGCTGACGATCTCTGCTCAAGAACTTAT 887  
Db 470 KNN 529  
QY 888 GTATCCCTCTGCTGTGCTGACAAACCTTCGATGGAACCTGATTCATCCATCC 947  
Db 530 NCKKNN 589  
QY 948 ATCATCTGCTGCTTTCGAAAATTCCTATGAGAGTGAGCCCTAGCCGTTCTCT 1007  
Db 590 NKNNN 649  
QY 1008 CAGTTTACTAGTCCATTTGTCAGTGCTTCGTAAGGCTTCCCCACATGTTGCTTC 1067  
Db 650 NNN 709  
QY 1068 AGTTATATGATGATGTTGACTGAGGGGCAAGTGTGACACCATCTGATGCTTTT 1127  
Db 710 KKTNNKKKKNN 769  
QY 1128 ACCGCTGTTACCAATTTCTTTTGT 1152  
Db 770 NNN 794

RESULT 7

CNS016LM



LOCUS CNS0161M 1101 bp DNA linear GSS 26-JUL-1999  
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC  
 BACN16J16 of DrosBAC library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 ACCESSION AL106910.1 GI:5624430  
 VERSION AL106910  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1101)  
 REFERENCE Genoscope.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the European Drosophila Genome Project (EDGP) -  
 http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC  
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
 d'Etude du Polymorphisme Humain) with funding provided by a MRC  
 project grant. The DNA was prepared from embryos by Alain Bucheton  
 and Genevieve Payan. It has been constructed in the vector  
 pBelBAC11.

FEATURES  
 source  
 1..1101  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACN16J16"  
 /clone\_lib="DrosBAC"  
 /plasmid="pBelBAC11"  
 /note="end : 77"  
 /note="end : 77"

BASE COUNT 222 a 80 c 146 g 113 t 540 others  
 ORIGIN

Query Match 3.5%: Score 41.8; DB 29; Length 1101;  
 Best Local Similarity 10.4%; Pred. No. 10; Mismatches 211; Indels 0; Gaps 0;  
 Matches 35; Conservative 91; Mismatches 211; Indels 0; Gaps 0;

QY 18 CACCAGCAATCTCTGGATTCTTCCGACGACGATGGATCCAGCTTCAGAGCAAA 77  
 DB 379 CACCAACTNN 438  
 QY 78 CACCAACAATCCAGATTGGACTTCATCCACAAGACACCTGGCCAGAGCCACA 137  
 DB 439 GGGNN 498  
 QY 138 GGTAGAGCTGGAGATTCCGACTGGGCTTACCCGACGAGAGGCTTTGGGGTG 197  
 DB 499 NNN 558  
 QY 198 GAGCCCTCAGGCTCAGGATACACAACCTTGGCAGATCCGCTCTGCTTCAC 257  
 DB 559 NNN 618  
 QY 258 CAATGCCAGTCAGAGAGGACCTTACCCGCTGCTCCACTTGGAGAAACATCTATC 317  
 DB 619 NTNTCTNTNN 678  
 QY 318 TCAAGCATGCACTGGAAGTCCACAACCTTCCAGCA 354  
 DB 679 MCMNN 715

RESULT 8  
 LOCUS AZ879026 308 bp DNA linear GSS 05-MAR-2001  
 DEFINITION RPCT-23-193P13.TU RPCT-23 Mus musculus genomic clone RPCT-23-193P13  
 ACCESSION AZ879026

VERSION AZ879026.1 GI:13197092  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 308)  
 REFERENCE Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatman,S., Akintet  
 B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.  
 and Fraser,C.M.  
 AUTHORS Mouse BAC End Sequences from Library RPCT-23  
 JOURNAL Unpublished  
 COMMENT Other GSSs: RPCT-23-193P13.TU  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCT-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC  
 Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
 Plate: 193 row: P column: 13  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source  
 1..308  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCT-23-193P13"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCT-23"  
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 28 a 130 c 35 g 115 t  
 ORIGIN

Query Match 3.5%: Score 41.6; DB 28; Length 308;  
 Best Local Similarity 49.5%; Pred. No. 6.8; Mismatches 109; Indels 0; Gaps 0;  
 Matches 107; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 572 TACCGCAGATCTAGACTGTGGTGACTTCTCAATTTCTAGGGGAATACCGTGT 631  
 DB 71 TTCTCAGAAATTAACATCGTTTGTATAGCTCTGAGTCGAGATCTTCTCAGGA 130  
 QY 632 GTCTGGCCAAATTCGAGATCCCACTCAATCATCATCAACATCTCTCTCCAA 691  
 DB 131 TTTTGACCTTATATACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190  
 QY 692 CTTCGCTGTTATGCTGATGTGTGCGGCGTTTATATATCTTCTCTTCATCTGCG 751  
 DB 191 CT 250  
 QY 752 TGTATGCTCATCTTCTTGTGTTCTTCTGAGCT 787  
 DB 251 TCCT 286

RESULT 9  
 LOCUS AZ413108 377 bp DNA linear GSS 03-OCT-2000  
 DEFINITION 1M0197017F Mouse 10kb plasmid UUCG1M library Mus musculus genomic

clone UUGC1M0197017 F, genomic survey sequence

ACCESSION

VERSION

## KEYWORDS

**SOURCE**  
**ORGANIS**

01000000

## REFERENCE AUTHORS

1 (bases 1 to 377)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Ratner, J., and T. J. ...

84374

### FILE

JOURNAL

Contact: Robert B. Weis  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
plate: 0197 row: 0 column: 17  
Seq primer: CGTGTAAACGACGCGCAGT  
Class: plasmid ends  
High quality sequence. stop: 377.

## FEATURES

BOUTE

/orgанизм="res\_nucleotid"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /sex="male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_id="Mouse 10kb plasmid UUGCM library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (<http://www.jax.org/resources/documents/dnares/>). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of PWD2 (gil4732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUN

ORIGIN

Query Match	3.5%	Score 41;	DB 28;	Length 377;
Best Local Similarity	56.2%	Pred No. 11;		
Matches 77; Conservative	0;	Mismatches 60;	Indels 0;	Gaps 0;

[illegible]

Oy	765	CTCTTGTGCTTCTTC	781
Db	324	CTTCTTCTTCTTCTTC	340

RESULT 10  
PH034701 /

BA034 / 01 /  
LOCIS

DEFINITION

**1**

**ACCESSION**

**VERSION**

## KEYWORDS

**SOURCE**

REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
1 (bases 1 to 745)

## AUTHORS

TITTLE

JOURNAL

**COMMENT**

Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel.: 301 838 0200  
Tel.: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPci-24. For BAC  
library availability, please contact Pieter de Jong  
(pejong@mail.nih.gov). Clones may be purchased from BACPAC  
Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end  
page: [http://www.tigr.org/cdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 267 row: K column: 17  
Seq primer: SP6  
Class: BAC ends.

## FEATURES

sons

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-267K1.7"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_idb="RPCI-24"
/note="Vector: pTRABAC1, Site_1: BamHI, Site_2: BamHI,
library was cloned in the pTRABAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

```

BASE COUN

ORIGIN

Query Match	3.5%;	Score 41;	DB 28;	length 745;
Best Local Similarity	53.4%;	Pred. No. 14;		
Matches	86;	Conservative	0;	Mismatches 75; Indels 0; Gaps 0;

QY	TTCCATATACACCAACCTCTGTCGCCAAGCTTGACCTGGATGCGATGTCCT	719
Db	CTCCTCTCCCTCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	544
QY	GCGGGCTTTATACATCTCTCATCTCGTCTGATGCTCATCTCTGTTGGTCT	779
Db	CTTCT	484
QY	TCTGACATATCAAGTATGCGCGGTTGCTCTAATTC	820
Db	TCTTCT	443

RESULT	11
BH063007/c	
LOCUS	
DEFINITION	BH063007 . . . . . 697 bp DNA linear GSS 18-JUL-2001
ACCESSION	RPCT-24-355LI9.TU RPCT-24 Mus musculus genomic clone RPCT-24-355LI9
VERSION	BH063007
KEYWORDS	BH063007.1 GI:14877191
SOURCE	GSS.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 697)
TITLE	Zhao,S., Nierman,W., Malek,J.J., Shatsman,S., Akinret,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shvartbeyan,A., Gebregorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
JOURNAL	Unpublshed Mus BAC End Sequences from Library RPCT-24
COMMENT	Other_GSSs: RPCT-24-355LI9..TV

FEATURES	Location/Qualifiers
SOURCE	1. .697

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-355L19"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_idb="RPCI-24"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;
RPCI-24 Mouse BAC Library produced by Peter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

```

	BASE COUNT	353 a	38 c	251 g	55 t	
Query Match						
Best Local Similarity	48.7%;					Score 40.8; DB 28; length 697;
Matches 111; Conservative	0;					Fred. No. 15; Mismatches 117; Indels 0; Gaps 0;
Qy	603	TCTCAATTTTGTAGGGGGGACATCCGCTGTGTCCTTGSCCAAAATTCAGATCCCAACTC				662
Db	598	TCTGTGTTCTTTGTGCTGTGCTCTTGCTGTACCTTTCTCTCTTTCTTCTTCTCTCTTC				539
Qy	663	CAATCACTACCAACCTTCCTGCTCCACACTTGTCTCGATTACGCTGGATGTGTGCG				722
Db	538	TTCTTCCTTCTCCTTCTCCTTCTCCTTCTCCTTCTCCTTCTCCTTCTCCTTCTTCTT				479
Qy	723	GCGTTTATCATCTTCTCTCTTCATCTCGCTGATAGCTCATCTTCTGTGGTTCTTCT				782
Db	478	CTTCT				419
Qy	783	GGACTATACAGGATGTGCCCGTTGCTGCTCTCATATCCAGAGATTC				830
Db	418	TGCTTTTACCTTTTCTCTTTCATTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCT				371

LOCUS	AV647972	629 bp	mRNA	linear	EST 15-JAN-2002
DEFINITION	AV647972 GLC Homo sapiens cDNA clone G1CBHD10 3', mRNA sequence.				
ACCESSION	AV647972				
VERSION	AV647972.1	GI:9868986			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Lin,F., Huang,Q., Cheng,Z., Li,N., Du,D., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.				
TITLE	Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding nonaneerous liver				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)				
MEDLINE	21625106				
PUBMED	11752456				
COMMENT	Contact: Zeguang Han				

```
/organism="Homo sapiens"  
/mol_type="mRNA"
```

```

BASE COUNT      178 a      159 c      122 g      169 t      1.others
ORIGIN
      /db_xref="taxon:9606"
      /clone="GLCBDH10"
      /tissue_type="corresponding non cancerous liver tissue"
      /dev_stage="Adult"
      /lab_host="SOLR"
      /clone_id="GLC"
      /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

```

Query Match	3.4%	Score 40.6	DB 9	Length 629
Best Local Similarity	52.7%	Pred. No. 17		
Matches	88	Conservative	0	Mismatches 79
				Indels 0
				Gaps 0
OY	1014	ACTAGTCCATTGTTGTCAGTGTTCGATGGAGGCTTTCCCCACATGTTGGCTTCAGTTAT	1073	
Db	10	ACTAGTACCAAGAGATACAGGCCCTATACATCTTTTCGATTATCATGACTATATTTCAT	69	
OY	1074	ATGGATGATGTTGTACTGGGGGCCAAGTCTGTACACCATTTGATGCCCTTTTAAACCGCT	1133	
Db	70	CTTGATGATATATCTGATTTGCTGCACAAATGAGAGCTGCAGTGTCCAAACCTTTTCAATGCT	129	
OY	1134	GTTAACCAATTTCTTTTGTCTTTGGGGATATACATTTAAACCCCTATATA	1180	
Db	130	ACCTCCACACTGCTTAAACCGTATAGGATATCGATACCACTTATCTCA	176	

RESULT 13			
LOCUS	AZ938251		
DEFINITION	A2938251 512 bp DNA linear GSS 26-APR-2001		
ACCESSION	2M0196F16R Mouse 10kb plasmid UDGCM library Mus musculus genomic		
VERSION	Clone UDGCM0196F16 R. genomic survey sequence.		
KEYWORDS	A2938251 A2938251.1 GI:13797782		
SOURCE	GSS.		
ORGANISM	Mus musculus (house mouse)		
	Mus musculus		

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 512)

Dunn, D., Aoyagi, A., Barber, M., Beacom, T., Duval, B., Hamil, C., ...

TITLE	Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL	plasmid inserts
COMMENT	unpublished
	Contact: Robert B. Weiss

```

Class: plasmid ends
High quality sequence stop: 512.
Location/Qualifiers
1..512
FEATURES
source
```

BASE COUNT	91 a	172 c	56 g	193 t
ORIGIN				

Query Match	3.4%	Score	40	DB	28	Length	512
Best Local Similarity	51.7%	Read	No.	22			
Matches	91	Conservative			85	Indels	0
						Gaps	0

QY 644 TTGGAGTCCCAACATCCAAATCAATCACCACACTCTGTGTCCTGCACTGTCTGTGTTA 704  
 Db 302 TTCTCTTCTCCCTTCT 362  
 QY 705 TGGCTGGATGTGTCTGGGGCTTTTATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 766  
 Db 362 TCT 422  
 QY 765 CTCTCTGTGGTCTCTCTGGAATCAAGATGTGCCGCTTGTCTCTAATTC 820  
 Db 422 TCT 477

## RESULT 14

LOCUS	DEFINITION	520 bp	DNA	linear	GSS 11-OCT-2002
B2128636	CH230-4-4009, TV CHORI-230 Segment 2 <i>Rattus norvegicus</i> genomic clone				
	CH230-4-4009, genomic survey sequence.				

SOURCE ORGANISM	Rattus norvegicus (Norway rat) Rattus norvegicus
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

REFERENCE 1 (bases 1 to 520)  
AUTHORS Zhao,S., Shetty,J.J., Shatman,S., Tescaye,G., Geer,K., Shwartsbeyn  
A., Gebregordais,E., Overton,L., Russell,D., Chen,D., Riggs,F., de  
Jong,P. and Fraser,C.M.  
TITLE Rat PAC Sequences from Library CHOR-230 MboI segment  
JOURNAL Unpublished  
COMMENT Other\_GSbs: CH230-44069.TU

Seq primer: T7  
Class: BAC ends.

BASE COUNT	22 a	257 c	8 g	233 e
ORIGIN				

Query Match	3.4%	Score 40	DB 28	length 520
Best Local Similarity	52.4%	Pred. No.	22	
Matches 88	Conservative	0	Mismatches 80	Indels 0
				Gaps 0

[illegible]

RESULT 15	CNS016B8/c	LOCUS	DEFINITION
CNS016B8	1101 bp	DNA	linear
Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15E23 of DrosBAC library from Drosophila melanogaster (fruit			GSS 26-JUL-1999

fly), genomic survey sequence.  
ACCESSION AL106526  
VERSION AL106526.1 GI:5622379  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
REFERENCE Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.  
FEATURES  
source Location/Qualifiers  
1..1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACN15823"  
/clone\_1b="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : SP6"  
BASE COUNT 207 a 62 c 154 g 259 t 419 others  
ORIGIN  
Query Match 3.4%; Score 40; DB 29; Length 1101;  
Best Local Similarity 23.2%; Pred. No. 30;  
Matches 82; Conservative 115; Mismatches 156; Indels 1; Gaps 1;  
QY 652 TCCCAACCTCAATCAGTCACTGACCACTCTGCTCCCACTTGTCTGTTATGCTGG 711  
DB 754 KIDYMTMTSNVMAAAMTACDHCANMKCDMMAMMASAGCTRTBCWMTCTNCTH 695  
QY 712 ATGTGTCGCGGCGTTTATCATCTTCCTCATCCTGCTGCTATGCTCATCTTCTTG 771  
DB 694 CTBCTTTTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 636  
QY 772 TTGCTTCTTCTGAGTCAAGTATGTTGCGGTTTGTCTTAATCCAGATCTTCA 831  
DB 635 MAYAMMGCTVNDMTMTMTSMMAHCAVSYCTVNTCTBNTYAHCTVMTNCKMHTCMM 576  
QY 832 ACCACGAGCAGGAGCATGACAGCTGACAGCTCTGCTCAAGAACTGTATGAT 891  
DB 575 AVCTCTCKCTCKBTCTMAAMAMAMAMAHAYBATCAVBTBTHTMTHTTMTCTBAT 516  
QY 892 CCTCTGCTGTGCTGTAACAACCTTCGATGGAATGCACTGTATTTCCATCCCATCA 951  
DB 515 MTAMCTCMKCAATATAMAAAAMNTVTVYGTAAHVSSGGAGGAGCRSCDCGDMGR 456  
QY 952 TCCTGGGCTTTGGAAATTTCTATGAGAGTGGGCTCAGCCGTTTCTCTGG 1005  
DB 455 WRCMTBATVTKTCTGCDCAHDDGGGVKGAARDGCGGAVAGRCRSTGTRSMTRG 402

Search completed: September 15, 2003, 02:58:16  
Job time : 1990 secs

**This Page Blank (uspto)**